

UNA PROPUESTA PARA INTRODUCIR EL CONCEPTO DE DISTRIBUCIÓN  
ESTADÍSTICA EN LA ESCUELA

CLAUDIA VIVIANA SANDOVAL VASQUEZ

UNIVERSIDAD PEDAGÓGICA NACIONAL  
FACULTAD DE CIENCIA Y TECNOLOGÍA  
DEPARTAMENTO DE MATEMÁTICAS  
BOGOTÁ  
2006

UNA PROPUESTA PARA INTRODUCIR EL CONCEPTO DE DISTRIBUCIÓN  
ESTADÍSTICA EN LA ESCUELA

CLAUDIA VIVIANA SANDOVAL VASQUEZ

Monografía

Felipe Fernández  
Asesor del trabajo

UNIVERSIDAD PEDAGÓGICA NACIONAL  
FACULTAD DE CIENCIA Y TECNOLOGÍA  
DEPARTAMENTO DE MATEMÁTICAS  
BOGOTÁ  
2006

# CONTENIDO

	Pág.
1. PRESENTACIÓN	1
1.1. Introducción	1
1.2. Objetivos Generales	2
1.3. Objetivos específicos	2
2. CONCEPTO DE DISTRIBUCIÓN ESTADÍSTICA	4
2.1. Introducción	4
2.2. Noción de variable en estadística	6
2.3. Distribuciones en el campo de la estadística descriptiva	8
2.3.1. Ejemplos	8
2.3.2. Distribuciones de frecuencia	9
2.3.3. Formas de representación de las distribuciones de frecuencia	12
2.4. Variable aleatoria	34
2.5. Distribuciones en el campo de la probabilidad	35
2.5.1. Clasificación de la probabilidad	35
2.5.2. Introducción distribuciones de probabilidad	37
2.5.3. Tipos de distribuciones de probabilidad	40
2.5.4. Ejemplos de distribuciones de probabilidad	40
3. ANÁLISIS DEL APRENDIZAJE	45
3.1. Introducción	45
3.2. Gráficas	50
3.3. Media aritmética	51

	Pág.
3.4. Variabilidad	53
3.5. Estadísticos de orden	53
3.6. Inferencia estadística	54
4. ANÁLISIS DE LA ENSEÑANZA	59
4.1. Introducción	59
4.2. Análisis de la colección alfa	61
4.2.1 Grado 6	61
4.2.2 Grado 7	71
4.2.3 Grado 8	77
4.2.4 Grado 10	78
4.2.5 Grado 11	80
4.3. Análisis de la colección Prentice hall	84
4.3.1. Grado 6	86
4.3.2. Grado 7	90
4.3.3. Grado 9	94
4.3.4. Grado 10	99
4.4. Resumen de temas considerados en las colecciones analizadas	102
4.5. Contextos	103
4.5.1. Biología	105
4.5.2. Física	106
4.5.3. Social	107
4.5.4. Política	107
4.6. Uso de nuevas tecnologías en el aula de clase	107
5. UNIDAD DIDÁCTICA	109
5.1. Ubicación en el currículo	109
5.2. Justificación	109

	Pág.
5.3. Prerrequisitos	111
5.4. Capacidades que desarrollarán los estudiantes	111
5.5. Contenidos	112
5.5.1. Conceptuales	112
5.5.2. Procedimentales	112
5.5.3. Actitudinales	112
5.6. Metodología	113
5.7. Actividades de primer tipo (Trabajo en el aula)	114
5.7.1. Tarea 1. Variables cuantitativas vs. variables cualitativas	114
5.7.2. Tarea 2. Variables en escala nominal y ordinal	116
5.7.3. Tarea 3. Variables en escala de intervalo y de razón	118
5.7.4. Tarea 4. Practica 1	119
5.7.5. Tarea 5. Variables continuas y discretas	120
5.7.6. Tarea 6. Estudios univariados y multivariados	121
5.7.7. Tarea 7. Recolección y organización de datos.	123
5.7.8. Tarea 8. Frecuencias absolutas	124
5.7.9. Tarea 9. Frecuencias relativas	126
5.7.10. Tarea 10. Frecuencia absoluta acumulada	128
5.7.11. Tarea 11. Frecuencia relativa acumulada	128
5.7.12. Tarea 12. Diagrama de barras vertical	130
5.7.13. Tarea 13. Diagrama de barras horizontal	131
5.7.14. Tarea 14. Conclusión diagrama de barras	131
5.7.15. Tarea 15. Practica 2	132
5.7.16. Tarea 16. Diagrama circular	133
5.7.17. Tarea 17. Practica 3	133
5.7.18. Tarea 18. Pictograma	134
5.7.19. Tarea 19. Practica 4	135
5.7.20. Tarea 20. Diagrama de tallo y hojas	136

	Pág.
5.7.21. Tarea 21. Análisis diagrama de tallo y hojas	139
5.7.22. Tarea 22. Análisis de datos	139
5.7.23. Tarea 23. Elaboración tabla. (Intervalos de clase)	140
5.7.24. Tarea 24. Practica 5	141
5.7.25. Tarea 25. Histograma (Estatura vs. número de alumnos)	142
5.7.26. Tarea 26. Polígono de frecuencias	143
5.7.27. Tarea 27. Análisis polígono de frecuencias	144
5.8. Actividades del segundo tipo basadas en Minitools	145
5.8.1. Manual introductorio de Minitools	146
5.8.2. Actividades	161
Conclusiones	171
Referencias Bibliograficas	173

## LISTA DE TABLAS

	Pág.
<b>Tabla 2.1.</b> Circulación de biblioteca pública	9
<b>Tabla 2.2.</b> Clan de los millonarios	10
<b>Tabla 2.3.</b> Resultados de examen aplicado a 110 estudiantes de bachillerato	10
<b>Tabla 2.4.</b> Distribución de frecuencias de los resultados de examen aplicado a 110 estudiantes de bachillerato seleccionados	11
<b>Tabla 2.5.</b> Distribución de frecuencias agrupadas basadas en los resultados de examen aplicado a 110 estudiantes de bachillerato	14
<b>Tabla 2.6.</b> Calificaciones de examen de 50 estudiantes	15
<b>Tabla 2.7.</b> Calificaciones de examen de 50 estudiantes	15
<b>Tabla 2.8.</b> Tabla de doble entrada	16
<b>Tabla 2.9.</b> Calificaciones de examen de 20 estudiantes	18
<b>Tabla 2.10.</b> Tabla de doble entrada	31
<b>Tabla 2.11.</b> Distribución de frecuencias de la edad de 100 personas	33
<b>Tabla 2.12.</b> Resultados de cierta prueba	38
<b>Tabla 2.13.</b> Probabilidad de los resultados de la prueba tabla 2.12.	38
<b>Tabla 4.1.</b> Trabajo sección dos. Alfa 6	64
<b>Tabla 4.2.</b> Trabajo sección dos. Alfa 6	65
<b>Tabla 4.3.</b> Ejercicio aplicación número 8. Alfa 6	66
<b>Tabla 4.4.</b> Lección 4.1 Alfa 6	68
<b>Tabla 4.5.</b> Lección 4.2 Alfa 6	69
<b>Tabla 4.6.</b> Alfa 7	72
<b>Tabla 4.7.</b> Lección 1.1. Alfa 7	75

	Pág.
<b>Tabla 4.8.</b> Lección 1.2. Alfa 7	75
<b>Tabla 4.9.</b> Sección 2.1. Prentice Hall 7	92
<b>Tabla 4.10.</b> Sección 2.2. Prentice Hall 7	92
<b>Tabla 4.11.</b> Sección 2.3. Prentice Hall 7	93
<b>Tabla 4.12.</b> Sección 4.1. Prentice Hall 9	97
<b>Tabla 4.13.</b> Resumen textos analizados	102
<b>Tabla 5.1.</b> Resumen tipos de variables	120
<b>Tabla 5.2.</b> Atentados contra torres de energía de todas las empresas	122
<b>Tabla 5.3.</b> Opiniones sobre la imagen de Uribe presidente de Colombia 2002 – 2006	122
<b>Tabla 5.4.</b> Opinión de los ciudadanos acerca de los postulados a la cámara	136
<b>Tabla 5.5.</b> Calificaciones examen de matemáticas	137
<b>Tabla 5.6.</b> Calificaciones examen de matemáticas	137
<b>Tabla 5.7.</b> Calificaciones examen de matemáticas organizados por intervalos de clase	140
<b>Tabla 5.8.</b> Efecto medicina “Tipo C” en 10 personas	163



## LISTA DE GRÁFICAS

	Pág.
<b>Gráfica 2.1.</b> Representación tallo y hoja datos tabla 2.1	19
<b>Gráfica 2.2.</b> Representación tallo y hoja organizado tabla 2.9	19
<b>Gráfica 2.3.</b> Representación doble tronco tabla 2.9	20
<b>Gráfica 2.4.</b> Histograma del número de "visitas" que ha tenido cierta página Web de acuerdo a la hora de la visita	21
<b>Gráfica 2.5.</b> Diagrama de barras que representa el comportamiento de las calificaciones parciales de tres alumnos de preparatoria	22
<b>Gráfica 2.6.</b> Diagrama de barras tridimensional que representa el porcentaje de PIB gastado en docencia e investigación por cinco países en el lapso de 1988 a 1999	23
<b>Gráfica 2.7.</b> Diagrama de barras horizontal que representa la población de un país ficticio	24
<b>Gráfica 2.8.</b> Gráfica de líneas que representa la población de un país ficticio	24
<b>Gráfica 2.9.</b> Polígono de frecuencias que representa el PIB gastado en docencia e investigación durante el año de 1990 en cinco países	26
<b>Gráfica 2.10.</b> Ojivas que representan el número de "visitas" que ha tenido cierta página Web de acuerdo a la hora de la visita	26
<b>Gráfica 2.11.</b> Ojiva que representa la frecuencia porcentual acumulativa de las "visitas" que ha tenido cierta página Web de acuerdo a la hora de la visita	27
<b>Gráfica 2.12.</b> Gráfica de áreas que representa la comparación del total de las especies de las familias del orden carnívoro y las que están amenazadas en México	28
<b>Gráfica 2.13.</b> Diagrama circular que representa la matrícula en licenciatura (en México) por áreas de conocimiento en el año de 1992	28

<b>Gráfica 2.14.</b> Diagrama circular que representa la matrícula en licenciatura (en México) por áreas de conocimiento en el año de 1992	29
<b>Gráfica 2.15.</b> Pictograma que representa la población de los estados unidos	30
<b>Gráfica 2.16.</b> Pictograma que representa la masa de tres planetas de nuestro sistema solar	30
<b>Gráfica 2.17.</b> Gráfico de dispersión de los datos de la tabla 2.10	32
<b>Gráfica 2.18.</b> Gráfico de burbujas que representa las diversas distancias de los planetas al sol	32
<b>Gráfica 2.19.</b> Diagrama de caja que representa los datos de la tabla 2.11.	33
<b>Gráfica 2.20.</b> Ejemplo particular distribución binomial	43
<b>Gráfica 2.21.</b> Representación función de probabilidad para la distribución normal.	44
<b>Gráfica 4.1.</b> Alfa 11	83
<b>Gráfica 5.1.</b> Reacciones de fallo acerca de la reelección del presidente Uribe electo año 2005	115
<b>Gráfica 5.2.</b> Ingresos operacionales de algunas páginas de Internet	115
<b>Gráfica 5.3.</b> Resultados encuestas para próximas elecciones	116
<b>Gráfica 5.4.</b> Resultados encuestas para próximas elecciones	117
<b>Gráfica 5.5.</b> Gráfica de estaturas de Juan y Pablo	118
<b>Gráfica 5.6.</b> Opiniones sobre las elecciones de presidente de Colombia 2006 – 2010	123
<b>Gráfica 5.7.</b> Cifra de venta de automóviles octubre 2005	130
<b>Gráfica 5.8.</b> Calificación acerca del presidente Uribe electo 2005	131
<b>Gráfica 5.9.</b> Situación de la bebida Coca cola en EE.UU. y México	132
<b>Gráfica 5.10.</b> Estado de la alimentación de los colombianos	133
<b>Gráfica 5.11.</b> Uso de celulares	134
<b>Gráfica 5.12.</b> Pictograma	135
<b>Gráfica 5.13.</b> Información de empresas y su valor en la bolsa	136
<b>Gráfica 5.14.</b> Diagrama de tallo y hoja calificaciones examen de matemáticas	138

	Pág.
<b>Gráfica 5.15.</b> Diagrama doble de tallo y hoja calificaciones examen de matemáticas	138
<b>Gráfica 5.16.</b> Histograma	142
<b>Gráfica 5.17.</b> Polígono de frecuencias. (Evolución del homicidio en Bogotá)	143
<b>Gráfica 5.18.</b> Polígono de frecuencias (Exportación de Banano)	144
<b>Gráfica 5.19.</b> Vista preliminar Minitools 1	147
<b>Gráfica 5.20.</b> Representación de frecuencias con barras de valores y selección de rango en Minitools 1	150
<b>Gráfica 5.21.</b> Representación de frecuencias con puntos en Minitool 2	151
<b>Gráfica 5.22.</b> Vista preliminar Minitools 2	152
<b>Gráfica 5.23.</b> Datos divididos en cuatro grupos	154
<b>Gráfica 5.24.</b> Datos divididos con diez frecuencias cada uno	154
<b>Gráfica 5.25.</b> Histograma con frecuencias de clase de 50 de ancho cada una	155
<b>Gráfica 5.26.</b> Caja de puntos	156
<b>Gráfica 5.27.</b> Vista preliminar Minitools 3	157
<b>Gráfica 5.28.</b> Four equal groups off 4 x equal groups	159
<b>Gráfica 5.29.</b> Two equal groups off x 9 equal groups	159
<b>Gráfica 5.30.</b> Cuadrícula de 6x6	160
<b>Gráfica 5.31.</b> Sistema de coordenadas	160
<b>Gráfica 5.32.</b> Representación de situación problema 1	162
<b>Gráfica 5.33.</b> Situación problema 1. Actividad 12	164
<b>Gráfica 5.34.</b> Representación de situación problema 2	166
<b>Gráfica 5.35.</b> Situación problema 2. Actividad 11	167
<b>Gráfica 5.36.</b> Representación de situación problema 3	168
<b>Gráfica 5.37.</b> Representación de situación problema 3. Four equal groups off 5 x equal groups	169
<b>Gráfica 5.38.</b> Representación de situación problema 3. Two equal groups off x 10 equal groups	170

## RESUMEN ANALÍTICO

**TIPO DE TRABAJO:** Tesis de Grado

**ACCESO AL DOCUMENTO:** Universidad Pedagógica Nacional

**TITULO DEL DOCUMENTO:** Propuesta para introducir el concepto de distribución estadística en la escuela.

**AUTORES:** SANDOVAL VASQUEZ, Claudia Viviana

**PUBLICACIÓN:** Bogota, 2006, 178 p.

**UNIDAD PATROCINANTE:** Universidad Pedagógica Nacional

**PALABRAS CLAVES:** distribución estadística, distribución de probabilidad, minitools, distribución de frecuencias, representación de distribuciones, análisis estadístico, variables estadísticas.

**DESCRIPCIÓN:** Este trabajo presenta una propuesta para introducir en la escuela la noción de distribución en el campo de la estadística. Para ello se hizo una revisión desde la perspectiva de las matemáticas escolares de dicho concepto, de sus representaciones y de asuntos relacionados con su enseñanza y aprendizaje. Dicha revisión contempla una mirada tanto al campo descriptivo y exploratorio como al inferencial.

**FUENTES:** Godino, J., Batanero C. & Cañizares M. (1996). *Azar y probabilidad, fundamentos didácticos y propuestas curriculares*. España: Editorial Síntesis.

**CONTENIDOS:**

1. Presentación
2. Concepto de distribución estadística
3. Análisis del aprendizaje
4. Análisis de la enseñanza
5. Unidad didáctica
6. Conclusiones
7. Referencias bibliográficas

**METODOLOGÍA:** El trabajo se basó en la realización de un análisis didáctico, en el que se consideró, entre otros asuntos: el estudio de la noción de distribución estadística, el análisis de una muestra de textos escolares y algunas cuestiones relacionadas con el aprendizaje y la enseñanza de la noción en cuestión. Finalmente, basada en dichos análisis, se llevó a cabo la elaboración de una propuesta didáctica.

**CONCLUSIONES:** La elaboración de esta monografía permite evidenciar no sólo la importancia de presentar situaciones reales o que parezcan reales en la enseñanza de la estadística sino también la existencia de una gran variedad de contextos en los que se encuentra presente dicha área, los cuales podrían ser utilizados para despertar el interés de los estudiantes. La propuesta hace evidente que es posible generar actividades en el aula de clase basadas en estudios que se llevan a cabo casi diariamente alrededor del educando.

Por otra parte, la indagación bibliográfica relacionada con el campo de investigación que atañe al tema de la noción de distribución estadística es bastante amplia y compleja. Debido a ello se debe reconocer que la elaboración de una propuesta didáctica en torno a toda la temática que se esboza en el marco conceptual fue demasiado ambiciosa, ya que a la luz del análisis de contenido debía dar cuenta de una parte descriptiva y de una parte probabilística.

Fecha de elaboración análisis: 15 de mayo de 2006

# CAPÍTULO 1

## PRESENTACIÓN

### **1.1 Introducción**

Hoy en día estamos en una era de información y tecnología, y por ello hay la necesidad de entender como se procesa la información, y como esta se puede llegar a traducir en conocimiento útil. Esta necesidad resalta la importancia de que los estudiantes entiendan los conceptos y los procesos que se llevan a cabo para el análisis y tratamiento de datos, ya que estos resultados son utilizados para realizar predicciones y tomar decisiones críticas y bien informadas. Para tal fin es imprescindible que se posea conocimiento de naturaleza estadística.

Pese a esto, en la escuela la estadística es tratada superficialmente, tal vez si sobran un par de horas al mes se utilizan para dar un tratamiento rápido a un par de conceptos, sin importar la claridad de estos, o en que medida puedan ser interesantes para los estudiantes.

Un ejemplo de este fenómeno se vio durante el desarrollo de la práctica en contextos educativos amplios, durante la cual hubo la oportunidad de observar la clase de estadística de dos de los niveles de educación, 6° y 8°, en los cuales se llevaba un manejo muy superficial sobre la estadística, se daban ciertos conceptos, luego ejemplos sobre estos y para terminar se proponían algunos ejercicios, en los cuales se utilizaban mecánicamente los resultados de los trabajos modelo que se habían llevado a cabo, sin observar así el trasfondo, la importancia y la utilidad con que cuenta la estadística.

Es necesario tener en cuenta que aunque a nivel popular se tiene una visión de la palabra estadística asociada solamente a recogida y representación de datos, promedios, índices, porcentajes, etc., la estadística moderna está mucho más relacionada con la interpretación de conjuntos de datos para ayudar a tomar decisiones, tarea que compete tanto a individuos de diferentes campos, por ejemplo de la medicina, la economía, la ingeniería y otros más,

como a grandes organizaciones compuestas por empresas o gobiernos, de esta forma actualmente la estadística es una ciencia básica, ya que resuelve problemas fundamentales y los conceptos que la integran se encuentran en la vida diaria.

Además por la prensa, radio y televisión, que son los medios actuales más utilizados por el hombre para adquirir información acerca del mundo exterior, se reciben noticias que obligan a las personas a tener idea clara sobre muchos conceptos estadísticos.

Con respecto a la educación escolar, en los lineamientos curriculares hay recomendaciones de que en los niveles de 5° a 8°, el conocimiento de naturaleza estadística no se puede restringir a sólo identificar el término medio de un conjunto de datos, a construir representaciones gráficas, o a entender determinados datos como respuestas a preguntas específicas y aisladas. Más bien se debe procurar que la enseñanza involucre al estudiante en un proceso donde la resolución de problemas esté inmersa en contextos que tengan sentido para éste.

Este proceso debe comprender desde la formulación de preguntas claves, la recogida, organización y representación de datos, el análisis de datos, la elaboración de conjeturas, y hasta la misma comunicación de la información.

Para tal fin es necesario identificar contextos de situaciones y problemas que motiven a los estudiantes. Por ejemplo para muchos de los educandos puede llegar a ser interesante el conocimiento estadístico ya que la publicidad, las predicciones, y la política social, se basan frecuentemente en el análisis de datos.

El propósito de esta monografía es el de aportar una propuesta que sirva de referencia para los maestros, en la cual ellos encuentren una guía de enseñanza del concepto de distribución estadística.

## **1.2 Objetivo general**

Realizar un análisis didáctico en torno al concepto de distribución estadística, y con base en éste, proponer una secuencia de actividades que respondan a intereses y motivaciones de los estudiantes para introducir tal concepto en la escuela.

## **1.3 Objetivos específicos**

En particular en este trabajo se quiere lograr los siguientes objetivos:

- Determinar elementos conceptuales que desde el punto de vista matemático y curricular, conforman el concepto de distribución.
- Identificar algunos de los errores y dificultades más comunes en torno al concepto en cuestión.
- Indagar sobre aproximaciones metodológicas de enseñanza en torno al concepto de distribución estadística.
- Determinar posibles contextos que sirvan como referencia para la elaboración de la unidad didáctica.
- Elaborar y presentar una secuencia de actividades, consonante con las indagaciones anteriores.



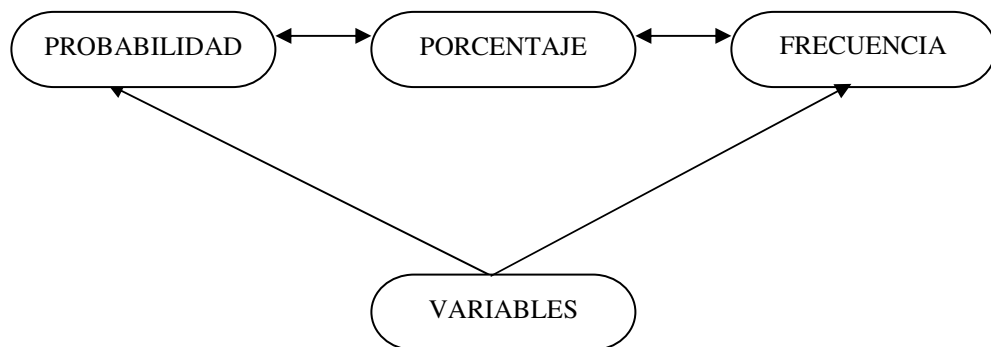
## CAPÍTULO 2

### CONCEPTO DE DISTRIBUCIÓN ESTADÍSTICA

#### 2.1. Introducción

En este capítulo se pretende presentar el concepto de distribución como una noción matemática compleja que tiene diversos significados, aplicaciones y/o interpretaciones según el campo de la estadística o el contexto de trabajo en el que se esté poniendo en juego.

De esta forma en primer lugar se dará una visión general sobre lo que se puede entender como una distribución en el campo de la estadística, usando como herramientas los conceptos fundamentales para el estudio de éste termino.



El concepto de distribución está íntimamente relacionado con cuatro nociones fundamentales: la idea de variable, la noción de frecuencia en la agrupación de datos, el concepto de proporcionalidad expresado a través de porcentajes o frecuencias relativas y la asignación de probabilidades asociadas a espacios muestrales, a través de la definición de variables aleatorias. La idea de variable debe considerarse tanto desde un enfoque inferencial a través de variable aleatorias, como desde un enfoque descriptivo a través de los diferentes niveles de medición (nominal, ordinal, intervalo y de razón). La noción de frecuencia toma especial importancia en el enfoque descriptivo en donde los procesos de recuento, organización y clasificación de datos son fundamentales. Los porcentajes

asociados a frecuencias permiten las comparaciones entre diferentes distribuciones, en particular para distribuciones con diferentes tamaños de muestras. Finalmente, la asignación de probabilidades desde las perspectivas clásica y bayesiana no tiene relación evidente con la idea de distribución que se deriva del enfoque descriptivo, pero si desde la visión frecuentista.

Desde el punto de vista de las aplicaciones que tiene la estadística se puede reconocer en la literatura disponible una división tradicional entre estadística descriptiva e inferencial (ver por ejemplo, Moya (1995)). Siguiendo esta división y para propósitos de concretar la idea de qué es una distribución se puede hacer una clasificación de este concepto en dos grupos: las distribuciones empíricas, y los modelos teóricos de distribución; el primero de ellos se considera de manera preponderante en la llamada estadística descriptiva, mientras que el segundo se encuentra en el campo de la teoría de la probabilidad y más precisamente en la inferencia estadística.

En la estadística descriptiva es usual hablar de una distribución de frecuencias como un listado de las frecuencias observadas de todos los resultados de una experiencia o de un experimento que se presentaron cuando éste sucedió; por otra parte, en la estadística inferencial una distribución de probabilidad es un modelo matemático de las probabilidades de todos los posibles resultados que podrían obtenerse si un experimento determinado se llevara a cabo.

Las distribuciones de probabilidad están relacionadas con las distribuciones de frecuencias. Una distribución de frecuencias teórica es una distribución de probabilidades que describe la forma en que se espera que varíen los resultados. Debido a que estas distribuciones tratan sobre expectativas de que algo suceda, resultan ser modelos útiles para hacer inferencias y para tomar decisiones en condiciones de incertidumbre.

Las distribuciones de probabilidad pueden basarse en consideraciones teóricas o en una estimación subjetiva de las posibilidades de eventos o también se pueden basar en la experiencia empírica u observada.

En este apartado queremos dar cuenta de cómo se concreta el concepto de distribución en cada uno de estos campos. Al final del capítulo se encuentra un cuadro con el cual se intenta hacer un esquema del resumen de ambos tipos de distribución.

## **2.2. Noción de variable en la estadística**

Un aspecto que resulta importante tocar para el estudio de las distribuciones es el concepto de variable. En el campo de la estadística una variable se define como una característica observable en un objeto de estudio, que puede tomar diferentes valores o expresarse en varias categorías.

Desde el punto de vista de su naturaleza, las variables se pueden clasificar en cualitativas y cuantitativas, las primeras son para referirse a elementos que tienen ciertos rasgos no medibles numéricamente, y las segundas, para aquellas que se pueden medir en una escala numérica.

Así mismo las variables cuantitativas se suelen dividir en variables discretas o variables continuas, siendo las primeras aquellas que están definidas sobre un número finito de valores o bien en un número infinito de valores pero numerables; mientras que las continuas son aquellas definidas sobre intervalos infinitos o acotados pero no numerables

Por otra parte, para clasificar los datos que resultan de procesos de medición o de recolección de información se utilizan diferentes tipos de escalas. Una de las clasificaciones más conocidas es: nominales, ordinales, de intervalo y de razón. El tipo de escala que se utilice depende de los objetivos del estudio y de la naturaleza de la variable.

La escala nominal consiste en clasificar los objetos o fenómenos según ciertas categorías, tipologías o nombres, dándoles una denominación o símbolo, sin que ésta implique alguna relación de orden, distancia o proporción entre los objetos, por ejemplo cuando un producto se rotula de acuerdo al cumplimiento de las especificaciones del diseño como conforme y no conforme, o como crítico, grave y menor, no se obtienen valores numéricos ni tampoco se pueden ordenar las observaciones, sólo tiene sentido la noción de clasificación.

Con la escala ordinal se establecen posiciones relativas de los objetos o fenómenos en estudio, respecto a alguna característica de interés, los numerales empleados en este tipo de escalas no son cuantitativos, sino que indican la posición en la serie ordenada. Un ejemplo pueden ser las opiniones de los clientes de un almacén que valoran la calidad de acuerdo a las siguientes respuestas, 1 excelente, 2 bueno, 3 regular, 4 malo y 5 pésimo.

La escala de intervalo representa un nivel de medición más preciso, además de que se establece un orden en las posiciones relativas de los objetos o individuos, también se mide la distancia entre los intervalos o las diferentes clases. Supongamos que se miden la temperatura de acero fundido, y se toman cuatro lecturas cada dos horas: 2050° F, 2100° F, 2150° F, 2200° F, estos datos pueden ser ordenados, además las diferencias entre los datos ordenados pueden ser comparadas, hay que tener en cuenta que en esta escala no hay un cero absoluto o real, el cero es arbitrario, por tanto no se puede decir que 76° F es el doble de 38° F.

Para terminar otro ámbito que resulta importante tocar en el tema de las distribuciones, es que los datos pueden combinarse simultáneamente. Así para un estudio o un análisis de datos se podrían considerar una, dos o más variables. Según la cantidad de variables involucradas se habla entonces de un estudio univariado, bivariado, o multivariado (Chou, 1972), y esto implica considerar distribuciones de cualquiera de estos tipos según el estudio.

### **2.3. Distribuciones en el campo de la estadística descriptiva**

En el desarrollo de un estudio estadístico, se obtienen diversos datos de observaciones hechas durante el mismo que necesitan ser organizados y clasificados para lograr una adecuada evaluación de la situación. El concepto de clasificación es familiar a cada una de las personas. Este concepto consiste en hacer un arreglo de datos, dentro de grupos creados o seleccionados de manera lógica, razonable y práctica teniendo en cuenta los propósitos de la investigación. De la clasificación de datos resulta lo que se conoce como distribución.

Esta clase de distribuciones se utilizan como una forma de resumir y clasificar información, de determinadas variables, o de características asociadas a muestras de datos. Según Wert (1938) las distribuciones se pueden clasificar en dos tipos: de valores o de frecuencias. En el primer caso se clasifican los objetos y las respectivas características que son recolectadas en la investigación y se pueden distinguir dos casos: de valores o eventos que son dispuestos de acuerdo al tiempo en que ocurren o de una ubicación geográfica. El segundo caso, se refiere al número de veces que se repite cierta característica, y su respectiva clasificación. A continuación se presentarán algunos ejemplos.

**2.3.1. Ejemplos** Un ejemplo de distribución donde los valores son clasificados de acuerdo al tiempo en el cual ocurren es el que se presenta en la tabla 2.1.

Un ejemplo de distribución donde los valores son clasificados de acuerdo al lugar en el cual ocurren es el que se presenta en la Tabla 2.2.

En la tabla 2.3 se tienen los datos recogidos sin organizar acerca de los resultados del examen del Coeficiente Intelectual (C.I.) de 110 estudiantes. Enseguida, en la tabla 2.4, se encuentran los datos organizados de mayor a menor, y cada uno de ellos con la respectiva frecuencia de los resultados del mismo examen. Esta última tabla se constituye en un ejemplo donde se encuentra una distribución de frecuencias.

*TABLA 2.1.*

*CIRCULACIÓN DE BIBLIOTECA PÚBLICA*

Año	Circulación	Año	Circulación
1920	976.607	1928	1.626.075
1921	1.018.436	1929	1.717.783
1922	1.067.324	1930	1.749.585
1923	1.021.672	1931	1.849.826
1924	1.075.658	1932	2.150.996
1925	1.109.438	1933	1.393.959
1926	1.262.171	1934	2.330.989
1927	1.427.188		

*Fuente: Reporte anual de la biblioteca nacional*

### **2.3.2. Distribuciones de frecuencia**

Este tipo de distribuciones tiene una buena diversidad de presentaciones y variaciones dependiendo de lo que se quiera describir o resaltar. Algunas de ellas se definen a continuación.

TABLA 2.2  
CLAN DE LOS BILLONARIOS

Personaje	Millones de dólares
Bill Gates	50.000
Warrent Buffet	42.000
Carlos Slim	30.000
Ingvard Kamprad	28.000
Lakshmi Mittal	25.000
Julio Mario Santodomingo	4.500
Luis Carlos Sarmiento	4.400

Fuente: Revista Semana. Del 13 de marzo al 20 de 2006. Página 17

TABLA 2.3  
RESULTADOS DE EXAMEN. DE 110 ESTUDIANTES DE BACHILLERATO SELECCIONADOS AL AZAR

154	131	122	100	110	119	121	128	112	93
133	1117	115	117	109	104	125	85	120	135
116	103	103	121	113	147	103	113	107	98
128	93	90	105	118	134	89	143	108	142
85	108	108	136	115	117	110	80	111	127
100	100	114	123	126	119	122	102	100	106
105	111	127	108	106	91	123	132	97	110
150	130	87	89	108	137	124	96	111	101
118	104	127	94	115	101	125	129	131	110
97	135	108	139	133	107	115	83	109	116
110	113	112	82	114	112	113	142	145	123

*Distribución de frecuencias absolutas.* Se refiere al número de veces que se repite cada valor o modalidad de la variable o atributo ( $n_i$ ), y debe satisfacer que  $\sum n_i = n$ , donde  $n$  es el número total de datos, individuos u observaciones de la población o muestra. Con este dato es posible hallar las tendencias del comportamiento de una población.

TABLA 2.4  
DISTRIBUCIÓN DE FRECUENCIAS DE LOS RESULTADOS DE CIERTO EXAMEN DE C.I.  
DE 110 ESTUDIANTES DE BACHILLERATO SELECCIONADOS AL AZAR

R	F	R	F	R	F	R	F	R	F	R	F	R	F		
154	1	145	1	136	1	127	3	118	2	109	2	100	4	91	1
153		144		135	2	126	1	117	2	108	6	99		90	1
152		143	1	134	1	125	2	116	2	107	2	98	1	89	2
151		142	2	133	2	124	1	115	4	106	2	97	2	87	1
150	1	141		132	1	123	3	114	2	105	2	96	1	86	
149		140		131	2	122	2	113	4	104	2	95		85	2
148		139	1	130	1	121	2	112	3	103	3	94	1	83	1
147	1	138		129	1	120	1	111	3	102	1	93	2	82	1
146		137	1	128	2	119	3	110	5	101	2	92		80	1

*Distribución de frecuencias relativas.* También se refiere a la frecuencia de ocurrencias la variable o atributo considerado, pero en términos relativos. Se suele representar como  $f_i$ , donde  $f_i = \frac{n_i}{n}$  y  $\sum f_i = 1$ . Este tipo de distribución hace evidentes los patrones existentes en un conjunto de datos.

*Distribución de frecuencias absolutas acumuladas.* Es la suma de las frecuencias absolutas correspondientes a los valores menores o anteriores, al valor sobre el cual se está evaluando dicha frecuencia, y se representa por medio de  $(N_i)$ .

*Distribución de frecuencias relativas acumuladas* Es la suma de las frecuencias relativas correspondientes a los valores menores o anteriores, al valor sobre el cual se esta evaluando dicha frecuencia, y se representa por medio de  $(F_i)$ .



**2.3.3 Formas de representación de las distribuciones de frecuencias** Hay varios tipos de representaciones para las distribuciones de frecuencias, según el tipo de distribución que se quiera representar y/o según el contexto de los datos.

***Representaciones tabulares univariadas:*** La representación tabular univariada, presenta la distribución de los datos de una variable mediante un cuadro o tabla, haciéndolos corresponder a cada uno de ellos su respectiva frecuencia. Si la diversidad del número de valores que asumen los datos es de gran magnitud, es necesario agrupar las observaciones en un número relativamente pequeño, estos grupos son llamados intervalos de clases, y se hacen con el fin de identificar los patrones presentes en la distribución. Las fronteras de una clase son llamadas límites de clase, y el promedio aritmético entre los límites superior e inferior recibe el nombre de punto medio de clase o marca de clase.

Los valores de los intervalos de clase deben ser mutuamente excluyentes, además el intervalo no debe ser tan amplio ni tan pequeño, para evitar perder la discriminación proporcionada por la medida original para no desvirtuar el objetivo que se busca con la agrupación.

Para generar las agrupaciones de los intervalos de clase no existe ninguna norma que pueda ser aplicada a todos los casos, pero normalmente se utilizan de 5 a 20 intervalos. En particular, Haber (1973) propone la siguiente técnica para hallar los intervalos: encontrar la diferencia entre el resultado más alto y más bajo de los datos originales, y añadir uno a esta para obtener el número total de resultados, o de resultados potenciales. Luego se divide la cifra entre 15 (el número potencial de intervalos de clase) para hallar el ancho del intervalo, si el número no es entero, los autores lo redondean al número impar más cercano, de manera que el número entero se encuentre siempre en la mitad del intervalo de clase, sin embargo no se comete ningún error si se redondea al número más cercano, a este valor lo llamaremos  $i$ .

Siguiendo se toma el dato más bajo, como el límite inferior del primer intervalo de clases y se agrega a este  $i-1$  para obtener el puntaje máximo del primer intervalo de clase. Por último el límite inferior de la clase siguiente, será el número entero consecutivo al máximo puntaje del intervalo de clase inferior.

Existen otras propuestas de fórmulas para estimar un número conveniente de intervalos de clase en Tukey (1977).

Hay diferentes recomendaciones que se suelen hacer cuando se elabora una tabla. Por ejemplo, cuando se presenta una tabla es necesario tener en cuenta que en el título de la misma indique a que estudio se refieren los datos, de donde provienen y cuando fueron tomados.

A manera de ejemplo se tomarán los datos de la Tabla 2.4, se distribuirán en intervalos de clase y se organizarán de nuevo en una tabla de 5 columnas (Tabla 2.5.), en la cual la primera columna esta compuesta por los intervalos de clase y en las cuatro siguientes se encuentran los diversos tipos de distribución de frecuencias que se han mencionado en un apartado anterior. Es posible tomar cada una de las frecuencias con su respectivo intervalo de clase en tablas diferentes, no es necesario que todos los tipos de frecuencia se encuentren en la misma tabla.

Como vemos en el ejemplo anterior, es posible que la suma de los valores correspondientes a las frecuencias relativas, no dé como resultado 1, este hecho se presenta porque al dividir cada una de las frecuencias entre el número total de datos, no da como resultado un número decimal con cifras finitas, así pues, se ha tomado una aproximación de tres cifras decimales lo cual generó cierta imprecisión en la frecuencias relativas.

Los intervalos de clase también pueden ser representados por medio de intervalos cerrados y abiertos, de la forma que se encuentra en la tabla 2.6.:

TABLA 2.5.

DISTRIBUCIÓN DE FRECUENCIAS AGRUPADAS DE LOS RESULTADOS DEL C.I. DE UN GRUPO DE ESTUDIANTES  
BASADAS EN LOS DATOS QUE APARECEN EN LA TABLA 2.4.

<i>Intervalo de Clase</i>	<i>Frecuencia Absoluta</i> $n_i$	<i>Frecuencia Relativa (<math>f_i</math>)</i>	<i>Frecuencia Absoluta Acumulada (<math>F_i</math>)</i>	<i>Frecuencia Relativa Acumulada (<math>N_i</math>)</i>
80 – 84	3	0.027	3	0.027
85 – 89	5	0.045	8	0.072
90 – 94	5	0.045	13	0.117
95 – 99	4	0.036	17	0.153
100 – 104	12	0.109	29	0.262
105 – 109	14	0.127	43	0.389
110 – 114	17	0.155	60	0.544
115 – 119	13	0.118	73	0.662
120 – 124	9	0.082	82	0.744
125 – 129	9	0.082	91	0.826
130 – 134	7	0.064	98	0.890
135 – 139	5	0.045	103	0.935
140 – 144	3	0.027	106	0.962
145 – 149	2	0.018	108	0.980
150 – 154	2	0.018	110	0.998
Total	110	0.998		

TABLA 2.6.  
CALIFICACIONES DE EXAMEN DE 50 ESTUDIANTES

<i>Intervalo de Clase</i>	<i>Frecuencia <math>n_i</math></i>	<i>Frecuencia Relativa (<math>f_i</math>)</i>	<i>Frecuencia Absoluta Acumulada (<math>F_i</math>)</i>	<i>Frecuencia Relativa Acumulada (<math>N_i</math>)</i>
[30 – 40)	4	0.08	4	0.08
[40 – 50)	6	0.12	10	0.20
[50 – 60)	8	0.16	18	0.36
[60 – 70)	12	0.24	30	0.60
[70 – 80)	9	0.18	39	0.78
[80 – 90)	7	0.14	46	0.92
[90 – 100)	4	0.08	50	1
<i>Total</i>	50	1		

En este caso la variable que se está trabajando es discreta, es por esta razón que dentro del intervalo sólo se van a tomar los valores enteros.

TABLA 2.7.  
CALIFICACIONES DE EXAMEN DE 50 ESTUDIANTES

<i>Intervalo de Clase</i>	<i>Frecuencia <math>n_i</math></i>	<i>Frecuencia Relativa (<math>f_i</math>)</i>	<i>Frecuencia Absoluta Acumulada (<math>F_i</math>)</i>	<i>Frecuencia Relativa Acumulada (<math>N_i</math>)</i>
30 – 39	4	0.08	4	0.08
40 – 59	14	0.28	18	0.36
60 – 69	12	0.24	30	0.60
70 - 89	16	0.32	46	0.92
90 – 100	4	0.08	50	1
<i>Total</i>	50	1		

También se pueden dar casos en donde los intervalos de clase no sean tomados con la misma amplitud, por ejemplo véase la Tabla 2.7.

En una tabla los datos también pueden ser presentados sin intervalos de clase. Esto aparece ejemplificado en la Tabla 2.5. Tales tipos de representación se utilizan cuando el tamaño de la muestra no es de gran magnitud.

**Tablas de doble entrada o divariadas:** Las tablas de doble entrada o bivariadas también son llamadas tablas de contingencias, ya que en ellas se representa la relación entre dos variables. Este tipo de tablas están formadas, en las cabeceras de las filas, por las categorías o valores de una de las variables, y en las cabeceras de las columnas, por los de la otra variable. En las casillas de la tabla, se indican las frecuencias o el número de elementos que reúnen a la vez a las dos categorías o valores de las dos variables que se cruzan en cada casilla.

Este tipo de tablas brindan información estadística de dos eventos relacionados entre sí, es útil en casos en los cuales los experimentos son dependientes de otro experimento. Es con base en tablas de doble entrada que se hacen los llamados análisis estadísticos bivariados.

TABLA 2.8

TABLA DE DOBLE ENTRADA

$T_1 / T_2$	Si	No
Si	13	18
No	26	32

**Representaciones gráficas:** Un diagrama es una especie de esquemático, formado por líneas, figuras, mapas, utilizado para representar, bien datos estadísticos a escala o según una cierta proporción, o bien los elementos de un sistema, las etapas de un proceso y las divisiones o subdivisiones de una clasificación.

Resulta importante aclarar que algunos autores toman como sinónimos el término gráfica y diagrama, como por ejemplo el caso Spiegel (1998), quien no hace ninguna salvedad entre la diferencia que puede existir entre los términos dentro de otro contexto.

Gran parte de la utilidad que tienen las representaciones gráficas en la estadística es la de proporcionar un medio para informar basado en los datos recopilados. La eficacia con que se pueda realizar tal proceso de información dependerá de la presentación de los datos. En general, la representación gráfica es una de las formas de representación más rápidas y eficientes para mostrar patrones o aspectos particulares de un grupo de datos, aunque también uno de los métodos de representación que más pueden ser manipulados o ser malinterpretados si no se tienen algunas precauciones básicas al elaborar gráficas.

Entre las funciones que cumplen los diagramas se pueden señalar las siguientes: Hacen más visibles los datos, sistemas y procesos, ponen de manifiesto sus variaciones y su evolución histórica o espacial, pueden evidenciar las relaciones entre los diversos elementos de un sistema o de un proceso y representar la correlación entre dos o más variables, sistematizan y sintetizan los datos, sistemas y procesos, aclaran y complementan las tablas y las exposiciones teóricas o cuantitativas, el estudio de su disposición y de las relaciones que muestran pueden sugerir hipótesis nuevas.

Algunas de las consideraciones que conviene tomar en cuenta al momento de realizar cualquier gráfica a fin de que la información sea transmitida de la manera más eficaz posible y sin distorsiones, son:

- El eje que represente a las frecuencias de las observaciones (comúnmente el eje vertical o de las ordenadas) debe comenzar en cero (0), de otra manera podría dar impresiones erróneas al comparar la altura, longitud o posición de las columnas, barras o líneas que representan las frecuencias.

- La longitud de los espacios que representan a cada dato o intervalo (clase) en la gráfica deben ser iguales.
- El tipo de gráfico debe coincidir por sus características con el tipo de información o el objetivo que se persigue al representarla, de otra manera la representación gráfica se convierte en un instrumento ineficaz, que produce más confusión que otra cosa, o se vuelve innecesario y/o productor de malas interpretaciones. Por ejemplo, si se desea representar la proporción de una población masculina en un país conviene más usar una gráfica de pastel o circular que una gráfica de barras al compararla contra la población femenina. Por un lado, se puede apreciar dicha proporción, y por el otro lado, se aprecia cuál de las dos poblaciones es mayor.

Existen también varios tipos de gráficas, o representaciones gráficas, utilizándose cada uno de ellos de acuerdo al tipo de información que se está usando y los objetivos que se persiguen al presentar la información, entre estas se encuentran principalmente, la representación de tronco y hojas, los histogramas, el diagrama de barras, la gráfica de líneas, polígono de frecuencias, ojiva, gráfica de áreas, diagrama circular, pictograma, gráfica de dispersión, gráfico de burbujas y diagrama de caja.

*Representación de tronco y hojas.* Es un método tabular que proporciona información rápida visual, para hacer una distribución de frecuencias. Esta representación se basa en la ordenación de los datos a manera de gráfico, pero utilizando la información numérica que aportan las decenas y las unidades.

La siguiente tabla que se presenta como ejemplo, contiene las calificaciones obtenidas en una prueba de matemáticas.

TABLA 2.9.  
CALIFICACIONES DE EXAMEN DE 20 ESTUDIANTES

78	93	61	100	70	83	88	74	97	72
66	73	76	81	83	64	91	70	77	86

Ahora se dispondrán cada uno de los datos presentados en la tabla separando las decenas de las unidades, es decir, las decenas se pondrán en una columna, en forma vertical, y las unidades a su derecha:

*GRAFICA 2.1.*  
*REPRESENTACIÓN TALLO Y HOJA DATOS TABLA 2.9*

6	1 6 4
7	8 0 4 2 3 6 0 7
8	3 8 1 3 6
9	3 7 1
10	0

Para aclarar un poco más el procesos de elaboración de este tipo de gráfico, es importante decir que el primer renglón por ejemplo, es decir el que presenta la configuración 6 | 1 6 4, indica que en la lista de datos se deben encontrar los valores 61, 66 y 64.

Como puede verse, este tipo de representación deriva su nombre del hecho de interpretar a cada renglón como la posición de un tronco y a cada dígito de la derecha como una hoja.

El procedimiento para realizarla es primero empezar con los troncos, es decir, con la columna de la izquierda, y después dato por dato ir llenando las hojas a la derecha de la línea vertical, en el tronco correspondiente.

Además, si se desean tener los datos ordenados, por facilidad de reconocer las frecuencias, se pueden ordenar las hojas en cada renglón para que la representación quede como sigue:

*GRAFICA 2.2.*  
*REPRESENTACIÓN TALLO Y HOJA ORGANIZADO TABLA 2.9*

6	1 4 6
7	0 0 2 3 4 6 7 8
8	1 3 3 6 8
9	1 3 7
10	0



En realidad una representación de tronco y hojas, suele presentar la misma información que la lista original de datos, pero de una manera mucho más compacta y manejable, especialmente si la lista de datos es muy grande.

Sin embargo, información más compleja resulta un poco más difícil de manejar, por lo que en ocasiones conviene redondear los datos, ignorar sus partes decimales o utilizar las centenas u otras posiciones de los números para las troncos. Para tales casos conviene hacer alguna anotación, o poner una nota, con el fin de que los lectores puedan identificar las adecuaciones realizadas y así poder interpretar lo que se quiere transmitir.

*Representación gráfica de doble tronco.* Para mostrar la información de manera más clara, es posible modificar el número de posiciones del tronco, aumentándola o disminuyéndola de acuerdo a las necesidades particulares de cada problema. Por ejemplo, con los datos del examen anterior, se pueden dividir en dos cada posición del tronco, utilizando la primera posición para disponer las hojas 0, 1, 2, 3 y 4, y la segunda posición para las hojas restantes.

GRAFICA 2.3.  
REPRESENTACIÓN DOBLE TRONCO TABLA 2.9

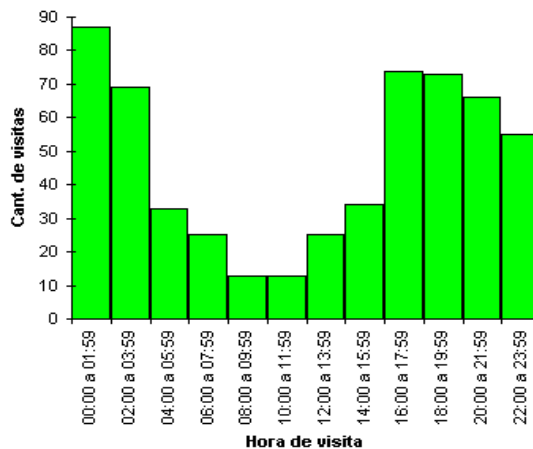
6-		1 4
6+		6
7-		0 0 2 3 4
7+		6 7 8
8-		1 3 3
8+		6 8
9-		1 3
9+		7
10-		0

Con esto se han duplicado el número de posiciones del tronco, con la intención de buscar un mayor detalle en la presentación o de identificar algún patrón que no se percibía muy claramente cuando los datos estaban agrupados de manera más compacta.

*Histograma.* Esta tipo de representación se obtiene al elaborar la gráfica de las frecuencias de los respectivos intervalos de clases de una representación tabular, en forma de rectángulos con sus alturas paralelas al eje vertical. En este caso, la frecuencia absoluta se indica en la escala del eje vertical y en el eje horizontal se señalan los límites de los intervalos de clases, o el punto medio de éstos. Por otra parte, la altura de cada barra se puede interpretar, como lo sugiere Yamane (1973), la densidad de la frecuencia de cada intervalo de clase. Conviene destacar que el concepto de densidad se refiere a una unidad base, por ejemplo refiriéndose a la densidad de una población, en el caso de un país como la India, ésta es alta ya que allí habitan 123 personas por kilómetro cuadrado, y en este caso la unidad base es el kilómetro cuadrado. En nuestro caso la unidad patrón es el intervalo de clase, y la frecuencia será el área de la barra, es decir la multiplicación de la densidad por la respectiva longitud del intervalo, es decir, como  $\frac{f}{\text{int}} * \text{int} = f$ . En este caso obtenemos que la densidad de la frecuencia es la misma frecuencia, ya que los intervalos son de la misma amplitud.

GRAFICA2.4.

HISTOGRAMA DEL NÚMERO DE "VISITAS" QUE HA TENIDO CIERTA PÁGINA WEB DE ACUERDO A LA HORA DE LA VISITA



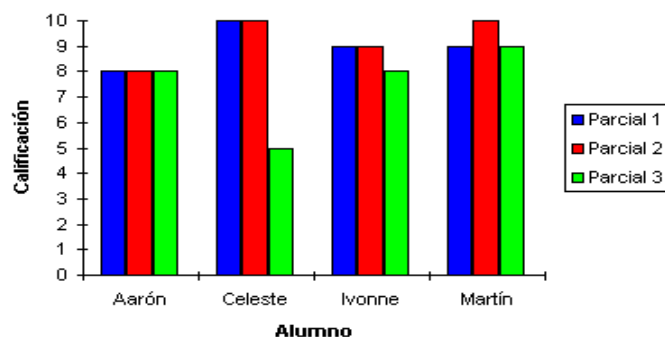
*Gráficos de barras.* Es un tipo de gráfico muy parecido al histograma que también se elabora con rectángulos y en donde se requiere que sus bases sean del mismo ancho y sus alturas equivalentes con las frecuencias. La diferencia con el histograma es que en este caso no es necesario tener una escala horizontal continua, por lo que las barras no tienen que aparecer juntas entre sí., ya que en el histograma se representan datos en escala continua y en este caso en discreta.

Otra observación pertinente es que se pueden representar en la misma gráfica los datos correspondientes a una misma variable producto de varias observaciones relativas a otra variable, utilizando las mismas escalas horizontales y verticales. Esto produce una gráfica de una observación multivariada, correspondiendo cada una de ellas a cada observación de una muestra o de una población. Es conveniente que cada serie de datos o de observaciones sea claramente distinguible de las demás.

El ejemplo que sigue muestra los resultados de las calificaciones en tres parciales diferentes, correspondientes a cuatro alumnos de cierto curso. Esta gráfica representaría más una serie de tiempo que una distribución de frecuencias, ya que, como variable se encuentra una calificación y no una frecuencia de calificaciones. Lo que hacen es una comparación de cada uno de los alumnos con respecto a los demás. Es interesante observar que la escala horizontal no es continua.

GRAFICA 2.5.

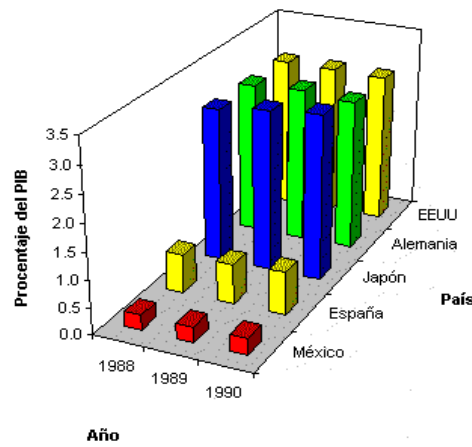
DIAGRAMA DE BARRAS QUE REPRESENTA EL COMPORTAMIENTO DE LAS CALIFICACIONES PARCIALES DE TRES ALUMNOS DE PREPARATORIA



Existe la posibilidad de representar gráficos compuestos de una manera "tridimensional", es decir, con gráficos que posean no sólo dos ejes, sino tres; y en los que los rectángulos son sustituidos por prismas de base rectangular. Un ejemplo es el siguiente:

GRAFICA 2.6.

DIAGRAMA DE BARRAS TRIDIMENSIONAL QUE REPRESENTA EL PORCENTAJE DE PIB GASTADO EN DOCENCIA E INVESTIGACIÓN POR CINCO PAISES EN EL LAPSO DE 1988 A 1999



FUENTE: REVISTA "CIENCIA Y DESARROLLO", 1994, XIX(114):12

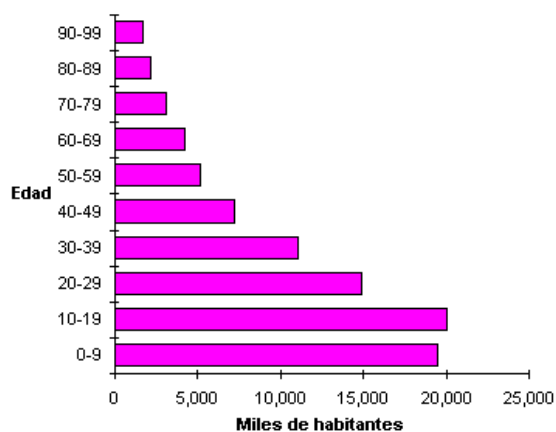
*Gráfico de barras horizontales.* Este tipo de gráfico se parece mucho a las gráficas de columnas, con la salvedad importante de que la función de los ejes se intercambia y el eje horizontal queda destinado a las frecuencias y el eje vertical a las clases.

Es muy común que este tipo de gráficos se utilicen para ilustrar el tamaño de una población dividida en estratos como, por ejemplo, son sus edades, es por esta razón que es utilizado para distribuciones bivariadas.

A este tipo de gráficos, en particular, se le llama pirámide de edades por su forma (ver la gráfica 2.7). Incluso, cuando se compara la población masculina y femenina por estratos de edades, se estila utiliza el lado izquierdo para la población de un sexo y el lado derecho para el otro, el resultado es una "pirámide" casi simétrica.

GRAFICA 2.7.

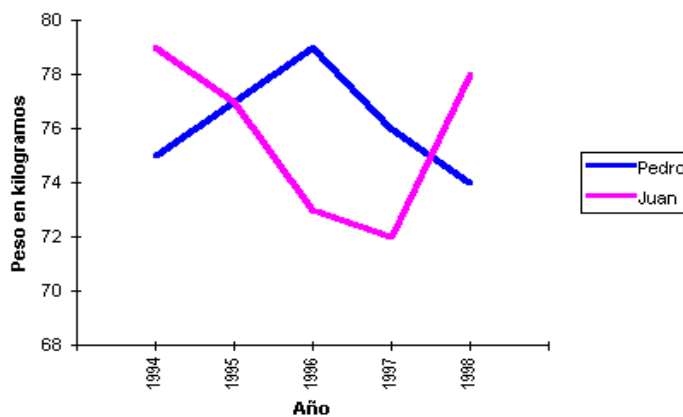
DIAGRAMA DE BARRAS HORIZONTAL QUE REPRESENTA LA POBLACIÓN DE UN PAIS FICTICIO



*Gráficas de líneas.* Se utiliza cuando los datos se relacionan entre sí, es decir, cuando podemos decir que existe cierta continuidad entre las observaciones. Tales casos son, por ejemplo, el crecimiento poblacional, la evolución del peso o estatura de una persona a través del tiempo, el desempeño académico de un estudiante a lo largo de su instrucción escolar o las variaciones presentadas en la medición realizada en algún experimento cada segundo o minuto. Las observaciones se representan por medio de una serie de puntos trazados en las intersecciones de las marcas de clase y las frecuencias de cada una, uniéndose consecutivamente con líneas.

GRAFICA 2.8.

GRÁFICA DE LINEAS QUE REPRESENTA LA POBLACIÓN DE UN PAIS FICTICIO



El ejemplo anterior muestra el comportamiento del peso corporal (en kilogramos) de dos individuos a lo largo de cinco observaciones anuales.

*Polígono de frecuencias.* Es posible transformar el histograma en otra forma empleada para las representaciones gráficas, el polígono de frecuencias, el cual se obtiene mediante la unión de los puntos medios de las barras con segmentos de rectas suavizadas.

La diferencia fundamental es que en el polígono de frecuencias se añaden dos clases con frecuencias cero: una, antes de la primera clase con datos, y la otra, después de la última clase con datos. El resultado es que se "sujeta" la línea por ambos extremos al eje horizontal y lo que podría ser una línea separada del eje se convierte, junto con éste, en un polígono.

Un ejemplo de polígono de frecuencias se encuentra ilustrado en la gráfica 2.9.

*Ojiva.* Es una gráfica similar al polígono de frecuencias, que representa las distribuciones de frecuencia acumuladas pero ésta se obtiene de aplicar parcialmente la misma técnica a la distribución, existen las ojivas mayor que y las ojivas menor que.

Existen dos diferencias fundamentales entre las ojivas y los polígonos de frecuencias:

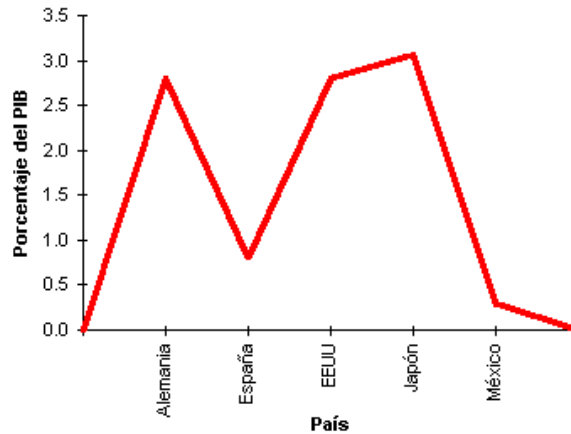
Un extremo de la ojiva no se "amarra" al eje horizontal, para la *ojiva mayor que* sucede con el extremo izquierdo; para la *ojiva menor que*, con el derecho.

En el eje horizontal en lugar de colocar las marcas de clase se colocan las fronteras de clase. Para el caso de la *ojiva mayor que* es la frontera menor; para la *ojiva menor que*, la mayor.

En la gráfica 2.10 se representan dos ejemplos de ojivas, a la izquierda la *mayor que*, a la derecha la *menor que*, utilizando los datos que se usaron para ejemplificar el histograma:

GRAFICA 2.9.

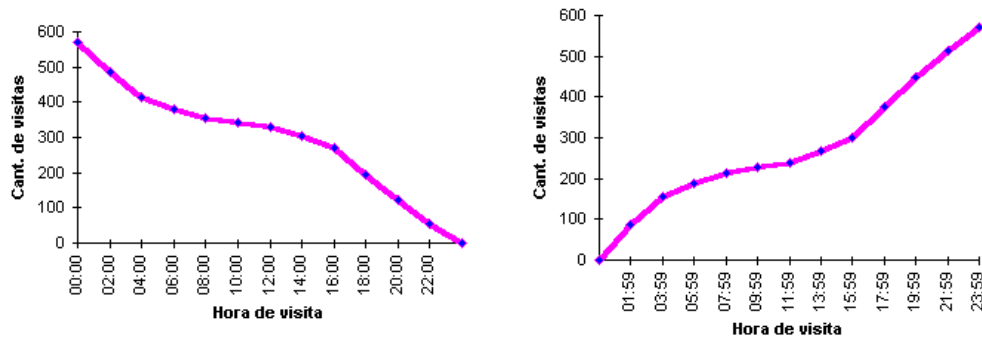
POLIGONO DE FRECUENCIAS QUE REPRESENTA EL PIB GASTADO EN DOCENCIA  
E INVESTIGACIÓN DURANTE EL AÑO DE 1990 EN CINCO PAÍSES



(FUENTE: REVISTA "CIENCIA Y DESARROLLO", 1994, XIX(114):12):

GRAFICA 2.10.

OJIVAS QUE REPRESENTAN EL NÚMERO DE "VISITAS" QUE HA TENIDO CIERTA PAGINA WEB DE ACUERDO A LA HORA  
DE LA VISITA



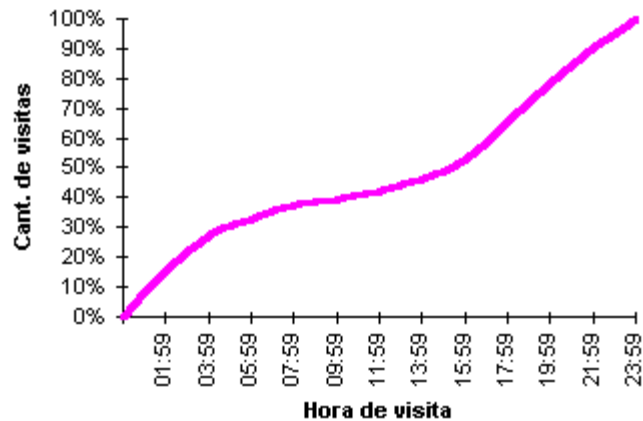
La ojiva mayor que (izquierda) se le denomina de esta manera porque viendo el punto que está sobre la frontera de clase "4:00" se ven las visitas que se realizaron en una hora *mayor* que las 4:00 horas (en cuestiones temporales se diría: *después de las 4:00 horas*). De forma análoga, en la ojiva menor que la frecuencia que se representa en cada frontera de clase son

el número de observaciones *menores que* la frontera señalada (en el caso de una variable como el tiempo sería el número de observaciones *antes* de la hora que señala la frontera).

Si se utiliza una distribución porcentual acumulativa entonces se obtiene una ojiva (*mayor que o menor que* según sea el caso) cuyo eje vertical tiene una escala que va del 0% al 100%. El siguiente ejemplo es la misma *ojiva menor que* que se acaba de usar, pero con una distribución porcentual:

GRAFICA 2.11.

OJIVA QUE REPRESENTAN LA FRECUENCIA PORCENTUAL ACUMULATIVA DE LAS "VISITAS"  
QUE HA TENIDO CIERTA PAGINA WEB DE ACUERDO A LA HORA DE LA VISITA



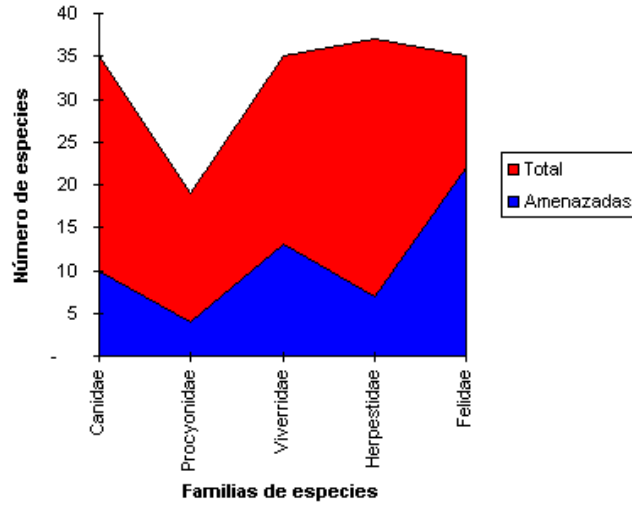
*Gráfica de áreas.* Este tipo de gráfica se utiliza en ocasiones para comparar datos bivariantes, esta consiste en rellenar el área que se encuentre debajo de las líneas que resultan de una gráfica de líneas. (Gráfica 2.12)

*Diagramas circulares.* Cuando lo que se desea es resaltar las proporciones que representan algunos subconjuntos con respecto al total, es decir, cuando se está usando una escala categórica, conviene utilizar una gráfica llamada de pastel o circular. Así pues esta gráfica se utiliza para representar las frecuencias relativas como porcentajes, donde 360° corresponde al 100%, a continuación se presenta un ejemplo de este tipo de gráfico. (Gráfica 2.13)



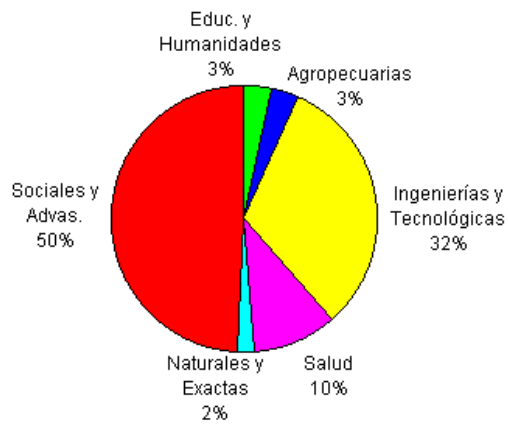
GRAFICA 2.12

GRAFICA DE ÁREAS QUE REPRESENTA LA COMPARACIÓN DEL TOTAL DE LAS ESPECIES DE LAS FAMILIAS DEL ORDEN CARNÍVORA Y LAS QUE ESTÁN AMENAZADAS EN MÉXICO



GRAFICA 2.13.

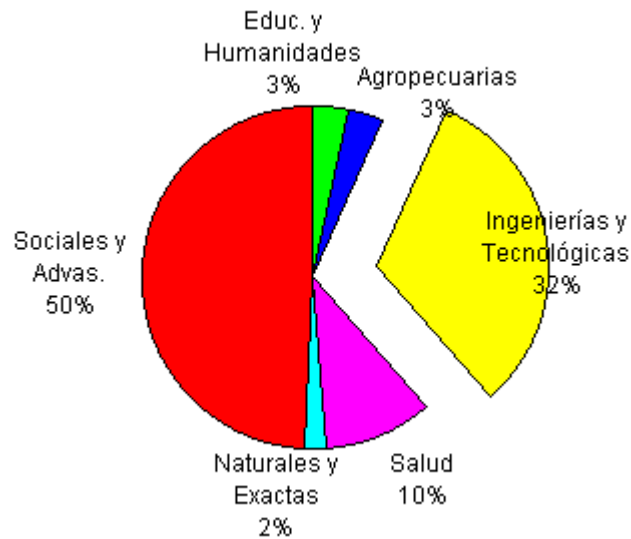
DIAGRAMA CIRCULAR QUE REPRESENTA LA MATRÍCULA EN LICENCIATURA (EN MÉXICO) POR ÁREAS DE CONOCIMIENTO EN EL AÑO DE 1992



Si se desea resaltar una de las categorías que se presentan, es válido tomar esa "rebanada" de la gráfica y separarla de las demás:

GRAFICA 2.14.

DIAGRAMA CIRCULAR QUE REPRESENTA LA MATRÍCULA EN LICENCIATURA (EN MÉXICO)  
POR ÁREAS DE CONOCIMIENTO EN EL AÑO DE 1992



Para utilizar este tipo de gráfica hay que tomar algunas precauciones. Por un lado, comparar dos gráficos circulares (por ejemplo, si se quisieran comparar las proporciones de matrículas en licenciatura por áreas de conocimiento en licenciatura para dos años distintos) resulta muy difícil y, por tanto, no es muy aconsejable.

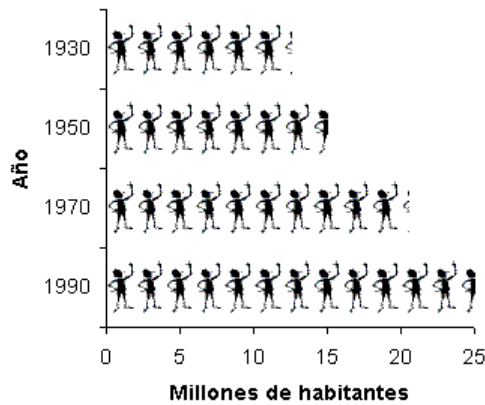
Por otro lado, en ocasiones existen categorías con pocas frecuencias (por ejemplo, dos o tres con frecuencias relativas menores al 1% cada una), haciendo que la gráfica resulte "pesada" y las etiquetas se encimen. Una posible solución es juntarlas en una sola categoría (por ejemplo, la típica "otras" o "varias"), pero entonces habría que ponderar si se hace una gráfica extra con dichas observaciones únicamente, haciendo la anotación pertinente, o simplemente si se ignoran por no resultar significativas.

*Pictograma.* Actualmente, y mucho en los medios masivos de comunicación, se utilizan gráficos para ilustrar los datos o los resultados de alguna investigación. Regularmente se utilizan dibujos para representar dicha información, y el tamaño o el número de estos

dibujos dentro de una gráfica queda determinado por la frecuencia correspondiente. A este tipo de gráfica se le llama pictograma y éstos son dos ejemplos (2.15, 2.16):

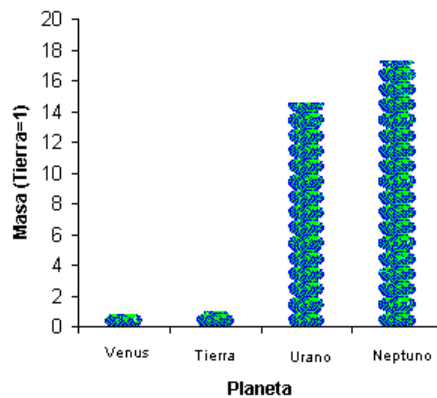
GRAFICA 2.15.

PICTOGRAMA QUE REPRESENTA LA POBLACIÓN DE LOS ESTADOS UNIDOS



GRAFICA 2.16.

PICTOGRAMA QUE REPRESENTA LA MASA DE TRES PLANETAS DE NUESTRO SISTEMA SOLAR



El primero representa la población de los Estados Unidos (cada hombrecillo representa a dos millones de habitantes), y el otro representa la masa de tres planetas de nuestro sistema solar tomando como unidad a la masa de la Tierra.

*Gráfica de dispersión.* Para trabajar los diagramas de dispersión, primero debemos saber qué es el análisis estadístico bivariado y cuáles son las características del mismo.

El análisis estadístico bivariado es un análisis que opera con datos referentes a dos variables y pretende describir y estudiar sus propiedades estadísticas. El análisis estadístico bivariable se orienta fundamentalmente a la normalización de los valores o frecuencias de los datos brutos, determina la existencia, dirección y grado de la variación conjunta entre las dos variables, lo que se realiza mediante el cálculo de los coeficientes de correlación pertinentes, calcula la covarianza o producto de las desviaciones de las dos variables en relación a sus medias respectivas y por último establece la naturaleza y forma de la asociación entre las dos variables en el caso de las variables de intervalo.

El diagrama de dispersión es un diagrama que representa gráficamente, en un espacio de abscisas y ordenadas, los puntos de dicho espacio que corresponden a los valores correlativos de una distribución bivariable conjunta, estos diagramas deben usarse cuando tenemos un análisis estadístico bivariable, es decir, de una tabla de datos que presente información de dos variables diferentes para una misma unidad de observación estadística. La ventaja que tiene este tipo de representación es que se puede graficar de una forma sencilla una distribución bivariable conjunta y la desventaja principal es que no funciona apropiadamente si sucede, por ejemplo, que una dupla se repite.

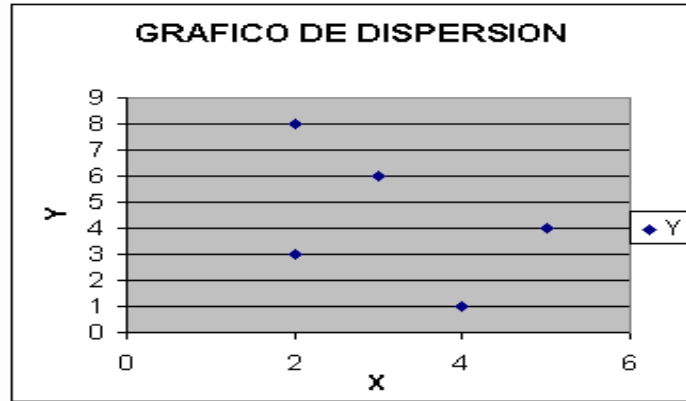
*TABLA 2.10*

*TABLA DE DOBLE ENTRADA*

	X	Y
A	2	3
B	4	1
C	5	4
D	3	6
E	2	8

GRAFICA 2.17

GRAFICO DE DISPERSION DE LOS DATOS DE LA TABLA 2.10

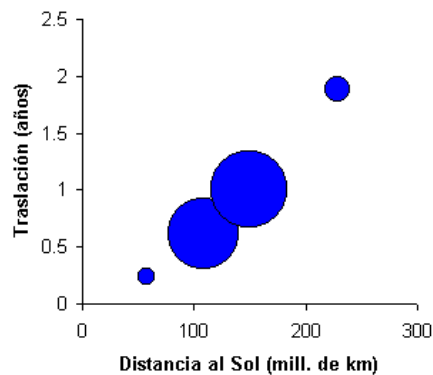


Uno de los usos de este tipo de gráficas es precisamente encontrar si las observaciones siguen algún patrón lineal o para detectar si existen valores atípicos.

*Gráficas de burbujas.* Es similar a las gráficas de dispersión, en estas se presenta la dispersión de las observaciones de la misma forma que aquéllas, pero se le añade la posibilidad de visualizar otra variable representada en el tamaño del punto, pues éstos se convierten en círculos (burbujas) con radios proporcionales a las magnitudes que representan.

GRAFICA 2.18

GRAFICA DE BURBUJAS QUE REPRESENTA LAS DIVERSAS DISTANCIAS DE LOS PLANETAS AL SOL



Este ejemplo compara la distancia que existe en cada uno de los planetas interiores de nuestro sistema solar al Sol contra el tiempo que necesitan para recorrer sus órbitas, y el tamaño de las burbujas indica la masa de cada planeta.

*Diagrama de caja.* Otra forma muy habitual, y muy útil, de resumir una variable de tipo numérico es utilizando el concepto de percentiles, mediante diagramas de cajas.

TABLA 2.11

DISTRIBUCIÓN DE FRECUENCIAS DE LA EDAD DE 100 PERSONAS

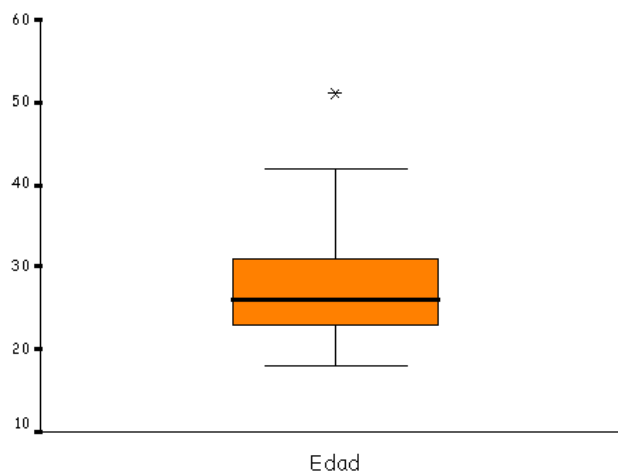
<i>EDAD</i>	18	19	20	21	22	23	24	25	26	27	28	29
<i>No</i>	1	3	4	7	5	8	10	8	9	6	6	4

<i>EDAD</i>	30	31	32	33	34	35	36	37	38	39	41
<i>No</i>	3	4	5	3	2	3	1	2	3	1	1

La gráfica 2.18 muestra un gráfico de cajas correspondiente a los datos de la tabla. La caja central indica el rango en el que se concentra el 50% central de los datos. Sus extremos son, por lo tanto, el 1<sup>er</sup> y 3<sup>er</sup> cuartil de la distribución. La línea central en la caja es la mediana. De este modo, si la variable es simétrica, dicha línea se encontrará en el centro de la caja. Los extremos de los "bigotes" que salen de la caja son los valores que delimitan el 95% central de los datos, aunque en ocasiones coinciden con los valores extremos de la distribución. Se suelen también representar aquellas observaciones que caen fuera de este rango (valores atípicos).

GRAFICA 2.19

DIAGRAMA DE CAJA QUE REPRESENTA LOS DATOS DE LA TABLA 2.11.



Esto resulta especialmente útil para comprobar, gráficamente, posibles errores en nuestros datos. En general, los diagramas de cajas resultan más apropiados para representar variables que presenten una gran desviación de la distribución normal.

Como se verá más adelante, resultan además de gran ayuda cuando se dispone de datos en distintos grupos de sujetos.

Antes de iniciar con las distribuciones en el campo de la probabilidad resulta conveniente definir una variable aleatoria.

## **2.4. Variable aleatoria**

Una variable es aleatoria si toma diferentes valores como resultado de un experimento aleatorio. Puede ser discreta o continua. Si puede tomar sólo un número limitado de valores, entonces es una variable aleatoria discreta. En el otro extremo, si puede tomar cualquier valor dentro de un intervalo dado, entonces se trata de una variable aleatoria continua.

En la teoría de la probabilidad una variable aleatoria se define como una función del espacio muestral en el que se encuentra definida una función de probabilidad sobre un conjunto numérico que puede ser el de los números reales o subconjunto de éste.

Se puede pensar en una variable aleatoria como un valor o una magnitud que cambia de una presentación a otra, sin seguir una secuencia predecible. Los valores de una variable aleatoria son los valores numéricos correspondientes a cada posible resultado de un experimento aleatorio.

La distribución de probabilidad de una variable aleatoria proporciona una probabilidad para cada valor posible, y estas probabilidades deben sumar 1.

## **2.5. Distribuciones en el campo de la probabilidad**

Luego de presentar las distribuciones en el campo de la estadística descriptiva, llega el momento de mencionarlas en el ámbito de la inferencia estadística, campo en el cual se busca explicar algunos de los fenómenos y se llegan a predecir otros, este hecho se logra utilizando como herramienta principal la teoría de la probabilidad. :

### **2.5.1. Clasificación de la probabilidad**

Existen tres maneras básicas de clasificar la probabilidad, éstas presentan fundamentos conceptuales bastante diferentes, el planteamiento clásico, el planteamiento de frecuencia relativa y el planteamiento subjetivo.

*Planteamiento clásico.* En éste se define la probabilidad de que un evento ocurra como el número de resultados en los que se presenta el evento dividido entre el número total de resultados posibles. Cada uno de los resultados posibles debe ser igualmente posible.

La probabilidad clásica, a menudo, se le conoce como probabilidad a priori, debido a que si se utilizan ejemplos previsible como monedas no alteradas, dados no cargados y mazos de barajas normales, entonces es posible establecer la respuesta de antemano, sin necesidad de lanzar una moneda, un dado o tomar una carta. No se tienen que efectuar experimentos para poder llegar a conclusiones.

Este planteamiento de la probabilidad tiene serios problemas cuando se intenta aplicarlo a los problemas de toma de decisiones menos previsible. El planteamiento clásico supone un mundo que no existe, supone que no existen situaciones que son bastante improbables pero que es posible concebir como reales. La probabilidad clásica supone también una especie de simetría en el mundo.

*Planteamiento frecuentista.* En el siglo XIX, los estadísticos británicos, interesados en la fundamentación teórica del cálculo del riesgo de pérdidas en las pólizas de seguros de vida y comerciales, empezaron a recoger datos sobre nacimientos y defunciones. En la



actualidad, a este planteamiento se le llama *frecuencia relativa de presentación* de un evento y define la probabilidad como la frecuencia relativa observada de un evento durante un gran número de intentos, o la fracción de veces que un evento se presenta a la larga, cuando las condiciones son estables.

Este método utiliza la frecuencia relativa de las presentaciones pasadas de un evento como una probabilidad. Determinamos qué tan frecuente ha sucedido algo en el pasado y esa cifra es usada para predecir la probabilidad de que suceda de nuevo en el futuro.

Cuando se utiliza el planteamiento de frecuencia relativa para establecer probabilidades, el número que se obtiene como probabilidad adquirirá mayor precisión a medida que aumentan las observaciones.

Una dificultad presente con este planteamiento es que la gente lo utiliza a menudo sin evaluar el número suficiente de resultados.

*Planteamiento Subjetivista.* Las probabilidades subjetivas están basadas en las creencias de las personas que efectúan la estimación de probabilidad. La probabilidad subjetiva se puede definir como la probabilidad asignada a un evento por parte de un individuo, basada en la evidencia que se tenga disponible. Esa evidencia puede presentarse en forma de frecuencia relativa de presentación de eventos pasados o puede tratarse simplemente de una creencia meditada.

Las valoraciones subjetivas de la probabilidad permiten una más amplia flexibilidad que los otros dos planteamientos. Los tomadores de decisiones pueden hacer uso de cualquier evidencia que tengan a mano y mezclarlas con los sentimientos personales sobre la situación.

Las asignaciones de probabilidad subjetiva se dan con más frecuencia cuando los eventos se presentan sólo una vez o un número muy reducido de veces.

Como casi todas las decisiones sociales y administrativas de alto nivel se refieren a situaciones específicas y únicas, los responsables de tomar decisiones hacen un uso considerable de la probabilidad subjetiva.

### **2.5.2. Introducción distribuciones de probabilidad**

Las distribuciones de probabilidad están relacionadas con las distribuciones de frecuencias. Una distribución de frecuencias teórica es una distribución de probabilidades que describe la forma en que se *espera* que varíen los resultados. Debido a que estas distribuciones tratan sobre expectativas de que algo suceda, resultan ser modelos útiles para hacer inferencias y para tomar decisiones en condiciones de incertidumbre.

Una distribución de frecuencias es un listado de las frecuencias observadas de todos los resultados de un experimento que se presentaron realmente cuando se efectuó el experimento, mientras que una distribución de probabilidad es un listado de las probabilidades de todos los posibles resultados que podrían obtenerse si el experimento se lleva a cabo.

Las distribuciones de probabilidad pueden basarse en consideraciones teóricas o en una estimación subjetiva de la posibilidad. Se pueden basar también en la experiencia.

De esta manera las distribuciones de probabilidad son clases especiales de distribuciones de frecuencia, sin embargo, las distribuciones de probabilidad se aplican a hechos no necesariamente observados y su valor numérico indica más un grado de posible ocurrencia que de ocurrencia observada. Las características de los eventos asociados a fenómenos naturales o sociales determinarán el tipo de distribución de probabilidad que se pueda aplicar para determinar un modelo del fenómeno y de esta forma explicar lo que podría ocurrir.

Si por ejemplo se tiene la siguiente distribución de frecuencia, acerca de las calificaciones de cierto examen:

TABLA 2.12

RESULTADOS DE CIERTA PRUEBA

Número de respuestas correctas en la prueba	Número de estudiantes
5	10
6	28
7	42
8	56
9	40
10	24
Total	200

Así pues hallamos la probabilidad de tener cierto número de respuestas correctas en la prueba, aplicando la noción de probabilidad frecuentista.

TABLA 2.13.

PROBABILIDAD DE LOS RESULTADOS DE LA PRUEBA TABLA 2.12.

Número de respuestas correctas en la prueba	Probabilidad
5	0.05
6	0.14
7	0.21
8	0.28
9	0.20
10	0.12
Total	1.00

Y como los eventos son mutuamente excluyentes, la suma de las probabilidades es igual a uno.

Pero la única tarea no consiste en hallar las probabilidades de las frecuencias, además es necesario estudiar el comportamiento de las distribuciones de frecuencias por medio de funciones de densidad de probabilidad, es tal vez por esta razón que son llamadas distribuciones teóricas de probabilidad.

Dependiendo del tipo de variable que se estudie, estas distribuciones de probabilidad se clasifican en dos grupos, las que estudian las variables aleatorias discretas y las que abordan el caso de las variables aleatorias continuas.

Se estudiarán con detalle algunas distribuciones específicas de probabilidad que han demostrado empíricamente ser modelos útiles para problemas prácticos, sin embargo presentan un carácter teórico ya que sus funciones de densidad o de probabilidad se deducen matemáticamente con base en ciertas hipótesis que se suponen válidas para los fenómenos aleatorios.

En general, los modelos de distribuciones de probabilidad se caracterizan no solamente a través de las relaciones analíticas o algebraicas de las variables involucradas, sino también por medio de valores no especificados, llamados parámetros, que sólo asumen un valor específico dependiendo del fenómeno particular que se esté considerando.

Se deben tener en cuenta que los parámetros pueden tomar cualquier valor dentro de un conjunto dado. Por ello, los parámetros más que definir una distribución, definen una familia de distribuciones de probabilidad, que tendrán características similares y la misma función genérica de probabilidad o de densidad de probabilidad.

Es necesario tener en cuenta las siguientes definiciones en el campo de las distribuciones de la probabilidad para trabajar con ellas:

### **2.5.3. Tipos de distribuciones de probabilidad**

Las distribuciones de probabilidad se clasifican como continuas y discretas. En la distribución de probabilidad discreta está permitido tomar sólo un número limitado de valores.

En una distribución de probabilidad continua, la variable que se está considerando puede tomar cualquier valor dentro de un intervalo dado. Las distribuciones continuas son una forma conveniente de presentar distribuciones discretas que tienen muchos resultados posibles, todos muy cercanos entre sí.

### **2.5.4. Ejemplos de distribuciones de probabilidad**

Como se trabajó anteriormente, existen distribuciones de probabilidad para variables aleatorias discretas y continuas. Para las primeras se trabaja con la función de probabilidad, esta función va a asignar probabilidades a cada uno de los valores de las variables aleatorias discretas.

La función de densidad de probabilidad es la que mide concentración de probabilidad alrededor de los valores de una variable aleatoria continua y la función de distribución es la que acumula probabilidades asociadas a una variable aleatoria.

Las distribuciones más estudiadas a nivel escolar, son la binomial en el caso discreto y la normal en el continuo, así pues estas son las que estudiaremos a continuación.

***Distribución binomial*** Esta es una de las distribuciones de probabilidad discreta más útiles, ya que sus áreas de aplicación incluyen inspección de calidad, ventas, mercadotecnia, medicina, investigación de opiniones y muchas otras más.

Este tipo de distribución se utiliza para los problemas que satisfacen los siguientes supuestos:

- El mismo experimento se efectúa  $n$  veces
- Cada ensayo tiene dos posibles resultados, usualmente referidos como éxito o fracaso; así, si  $p$  representa la probabilidad de éxito de un ensayo, y  $q$  la probabilidad de fracaso, se tiene que,  $p + q = 1$  o  $q = 1 - p$ .
- El resultado de cada ensayo es independiente del resultado de cualquier otro ensayo y el valor  $p$  de la probabilidad de éxito permanece invariable durante la replicación de los  $n$  experimentos.

*Representación analítica.* Sea  $X$  una variable aleatoria que representa el número de éxitos en  $n$  ensayos, y  $p$  la probabilidad de éxito con cualquiera de éstos. Se dice entonces que  $X$  tiene una distribución binomial con función de probabilidad dada por:

$$p(x; n, p) = \begin{cases} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} & x = 0, 1, 2, \dots, n \\ 0 & \text{para cualquier otro valor} \end{cases}$$

Para deducir de donde sale esta distribución primero se encuentra la probabilidad de obtener en  $n$  ensayos,  $x$  éxitos consecutivos seguidos de  $n-x$  fracasos consecutivos. Por hipótesis los  $n$  ensayos son independientes, entonces:

$$\underbrace{p \times p \times \dots \times p}_{\{x \text{ términos}\}} \times \underbrace{(1-p) \times (1-p) \times \dots \times (1-p)}_{\{(n-x) \text{ términos}\}} = p^x (1-p)^{n-x}$$

Lo cual se cumple en cualquier orden, así pues la probabilidad de tener  $x$  éxitos y  $n-x$  fracasos en cualquier orden es el producto de  $p^x (1-p)^{n-x}$  por el número de órdenes distintos, éste es el número de combinaciones de  $n$  objetos tomando  $x$  a la vez.

Recibe este nombre ya que tiene una evidente conexión con el binomio de Newton, es decir, con la expansión de la forma  $(x + y)^n$ . Por un lado sabemos que:

$$(x + y)^n = x^n + {}_n C_{n-1} x^{n-1} y + {}_n C_{n-2} x^{n-2} y^2 + \dots + {}_n C_1 x y^{n-1} + y^n$$

Donde  ${}_n C_x$  representa el valor de los coeficientes binomiales. Por ejemplo, para  $n = 3$  se tiene:

$$(x + y)^3 = x^3 + {}_3 C_2 x^2 y + {}_3 C_1 x y^2 + y^3 = x^3 + 3x^2 y + 2x y^2 + y^3$$

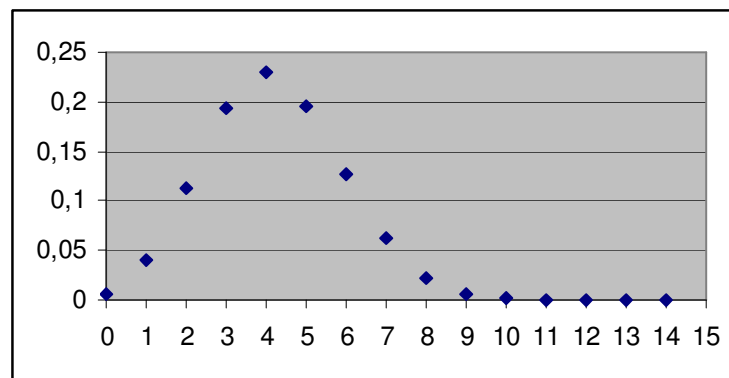
En la distribución binomial las probabilidades son términos del teorema binomial si se reemplazan a  $x$  por  $p$ , y a  $y$  por  $q$ :

$$(p + q)^n = p^n + {}_n C_{n-1} p^{n-1} q + {}_n C_{n-2} p^{n-2} q^2 + \dots + {}_n C_1 p q^{n-1} + q^n$$

*Representación gráfica.* Para mostrar un ejemplo de la gráfica de una distribución binomial daremos los siguientes valores a los parámetros:  $n = 15$  y  $p = 0.3$ .

GRÁFICA 2.20.

EJEMPLO PARTICULAR DISTRIBUCIÓN BINOMIAL



**Distribución continua normal** Esta distribución es de gran importancia dentro de la teoría de la probabilidad, puesto que las distribuciones de muchas estadísticas muestrales convergen hacia la distribución normal conforme crece el tamaño de la muestra. Por esta razón es una distribución fundamental en la inferencia estadística. En la literatura disponible no se encuentra de donde surge dicha distribución, pero es importante saber que la distribución de promedios converge a la distribución normal.

*Representación analítica.* Se dice que una variable aleatoria X se encuentra normalmente distribuida si su función de densidad de probabilidad esta dada por:

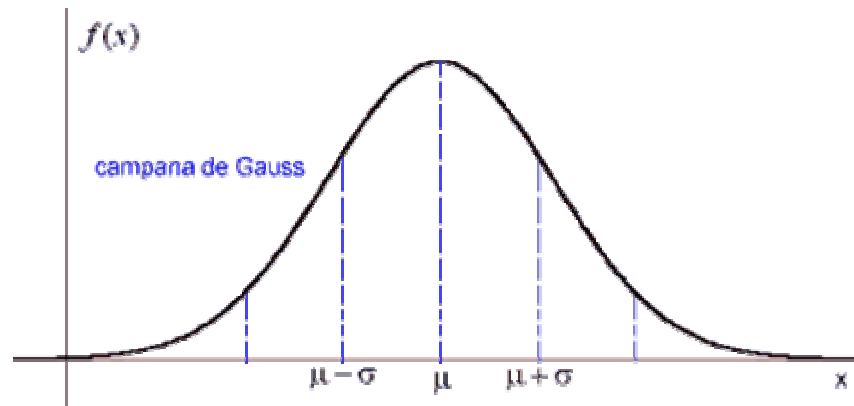
$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]. \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$$

Los parámetros de la distribución normal son  $\mu$  y  $\sigma$  los cuales son la media y la desviación estándar de X, respectivamente, para cualquier par de valores la gráfica es simétrica y tiene forma de campana. Si se obtienen las dos primeras derivadas de  $f(x; \mu, \sigma)$  con respecto a x y se igualan a 0, se tiene que el valor máximo de  $f(x; \mu, \sigma)$  ocurre cuando  $x = \mu$ , y los valores  $x = \mu \pm \sigma$  son las abscisas de los dos puntos de inflexión de la curva.

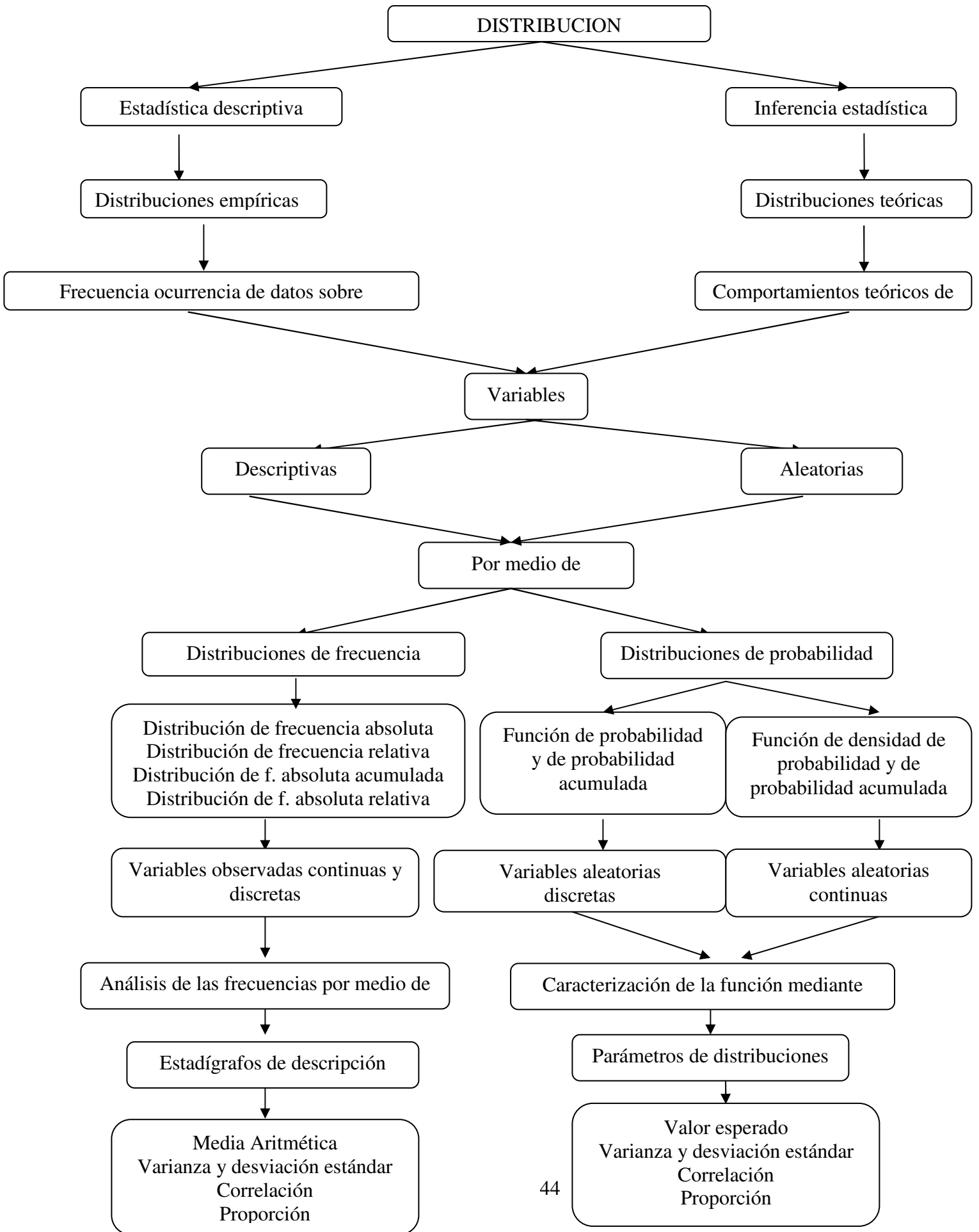
*Representación gráfica*

GRAFICA 2.21.

REPRESENTACIÓN FUNCIÓN DE PROBABILIDAD PARA LA DISTRIBUCIÓN NORMAL.







## CAPITULO 3

### ANÁLISIS DEL APRENDIZAJE

El aprendizaje de la estadística ha tomado mucha importancia a través de los años, esto se debe a la trascendencia que ésta ha venido adquiriendo en el diario vivir de una persona; desde en un día de ocio, hasta en uno de trabajo, va a estar presente no sólo la estadística, también la probabilidad.

Desafortunadamente, la estadística había recibido, hasta hace poco tiempo atrás, menos atención que otras ramas de las matemáticas; por esta razón la mayor parte de la investigación se ha llevado a cabo en situaciones experimentales, en lugar de situaciones escolares, además muchos estudios se centran en niños muy pequeños o en estudiantes de universidad, siendo escasa la investigación en las edades de 11 a 16 años. Por otra parte, las primeras investigaciones en el campo han sido efectuadas por psicólogos, donde estas deberían ser llevadas a cabo por parte de educadores matemáticos. Actualmente la situación está cambiando y esto se nota por ejemplo en trabajos como los de Batanero y demás profesores del grupo de investigación en educación estadística de la Universidad de Granada en España<sup>1</sup>. La importancia de dicha área se puede ver desde varios puntos, ya que por ejemplo desde la perspectiva de la psicología cognitiva, los resultados de investigaciones sobre educación estadística sugieren la posibilidad de generar estructuras cognitivas más competentes.

A continuación haremos un recuento de algunos de los resultados, que al respecto, llaman la atención.

Nisbett, Krantz, Jepron y Kunda (1983) estudiaron el razonamiento inferencial de cuatro grupos de personas con distintos niveles de formación estadística. Los sujetos eran

---

<sup>1</sup> Visítese, por ejemplo, la página <http://www.ugr.es/~batanero/> para tener una idea de la producción investigativa en este campo.

estudiantes de colegio (unos con formación estadística y otros no), graduados y doctores con varios años de experiencia. Se les presentó una situación problemática y los sujetos debían explicar por escrito lo que allí ocurría. Los resultados muestran que las personas sin preparación estadística no dan respuestas de tipo estadístico mientras que un amplio grupo de personas (80%) con preparación en estadística utilizan estos datos. Sobre el mismo tema Fong, Krantz y Nisbett (1986) realizaron cuatro experimentos en los que se pretendía estudiar el efecto de la enseñanza en el razonamiento estadístico. Se evidenció que la enseñanza estadística influye en el razonamiento estadístico de las personas sobre problemas cotidianos. Sus resultados se pueden resumir de la siguiente manera:

- El efecto de la enseñanza se produce con independencia del método, del contexto y del tipo de sujeto.
- La enseñanza provee beneficios con independencia de su duración.
- El efecto positivo de la enseñanza se produjo no solamente cuando el contexto de las pruebas de evaluación de resultados coincidía con el contexto en el que fueron enseñados, sino también cuando no fue así.
- La mejora en el pensamiento estadístico sucedió tanto en alumnos universitarios como también en alumnos de instituto y en adultos.
- Tanto la frecuencia como la calidad del pensamiento estadístico acaeció no sólo con eventos que normalmente están asociados con la incertidumbre y la probabilidad, sino también, para los que raramente se asocian con tales conceptos.

Es indudable que las estrategias estadísticas se emplean más en contextos conocidos, es así que Nisbett, Krantz, Jepron y Kunda (1983) han encontrado que el uso del razonamiento estadístico está asociado, incluso en personas sin formación estadística, con la familiaridad relativa al conocimiento del contenido del que se realiza la inferencia.

Para Castañeda y Rodrigo (1993), “en los dominios más conocidos, el espacio muestral o distribución de los sucesos, están más claros y por tanto las personas utilizan este dato. En

cambio en los dominios subjetivos los sujetos tienden a utilizar menos estrategias estadísticas”. Estos autores realizaron un estudio con 40 profesores universitarios de matemáticas y física, encontrando que en contextos cotidianos no se utilizaban estrategias de tipo estadístico.

La investigación sobre la naturaleza de las concepciones que poseen los estudiantes en estadística ha mostrado que los estudiantes tienen ideas acerca de la probabilidad y de la estadística que están reñidas con la teoría aceptada. *“La existencia de estas ideas fuertemente arraigadas pueden explicar en parte por qué el aprendizaje de la probabilidad y la estadística es especialmente problemático”* (Konold (1995)). Este autor resume los resultados de sus numerosas investigaciones resaltando tres conclusiones que tienen importantes implicaciones para la evaluación de las comprensiones intuitivas de los estudiantes: a) los alumnos llevan a clase intuiciones básicamente incorrectas que se encuentran muy arraigadas; b) Estas intuiciones son difíciles de cambiar; y c) alterarlas es complicado por el hecho de que un estudiante lleva arraigadas múltiples y a menudo contradictorias creencias sobre una situación particular.

Pero no sólo Konold aporta evidencias para aseverar las afirmaciones anteriores. Garfield y delMas (citado en Shaughnessy (1992)) examinaron, con un programa de ordenador desarrollado por ellos (Coin Toss), los efectos de una experiencia de ordenador con estudiantes sin conocimientos estadísticos sobre sus concepciones erróneas acerca de conceptos como el de variabilidad muestral, efectos del tamaño de la muestra en la variabilidad, la independencia y la aleatoriedad. Sus resultados muestran que un gran número de alumnos persisten en las concepciones erróneas sobre el tamaño de la muestra y la variabilidad después de la experiencia. Estos resultados concuerdan con los de Well et al. (1990) quienes encontraron que muchos estudiantes seguían sin comprender el efecto del tamaño de la muestra en la variabilidad de la media después de realizar varias experiencias de modelización de distribuciones muestrales con el ordenador. Konold (1995) no puede hacer recomendaciones explícitas del tipo de instrucción que ha demostrado ser efectiva para alterar las concepciones referidas anteriormente. De modo que daremos un vistazo a

algunas investigaciones que pueden orientar sobre cual es la naturaleza del razonamiento estadístico.

El razonamiento estadístico se puede definir como la manera en la que los sujetos razonan con ideas estadísticas o dan sentido a la información estadística (Garfield (1998)). Esto abarca las interpretaciones de conjuntos de datos, representaciones de datos, o resúmenes estadísticos de datos. Implícito al razonamiento estadístico está la comprensión de importantes ideas como distribución, incertidumbre, aleatoriedad y muestreo.

Muchas investigaciones tanto en el campo teórico como en el experimental, se refieren solo al error que comete el estudiante cuando hace cierto ejercicio. Esta manera de analizar, no proporciona resultados útiles, ya que el estudiante puede estar distraído, o simplemente se puede afirmar que la tarea es muy difícil para él, es necesario saber que los errores no son imprevisibles ya que es posible hallar regularidades, ya sea de acuerdo a las tareas propuestas, a las circunstancias o hasta con el individuo que se esta evaluando. El fin principal de la investigación en didáctica es hallar dichas constantes, y de esta forma llegar a construir modelos.

Algunos autores, como Radatz (1980), consideran el análisis de errores como “una estrategia de investigación prometedora para clarificar cuestiones fundamentales del aprendizaje matemático”. Un principio que es asumido en psicología educativa es el enunciado por Ausubel y cols. (1983): “el factor más importante que influye en el aprendizaje es lo que el alumno ya sabe. Averígüese esto y enséñese consecuentemente”. El interés reciente de los estudios de didáctica por las concepciones de los estudiantes (Confrey, 1990) sería una consecuencia de este principio. El problema que se puede presentar en didáctica, es la limitación que pueden llegar a tener estas cuestiones, por el hecho de que en algún momento no podrán ser aplicadas en términos más generales, así pues se tiene la existencia de errores y dificultades especiales, a las cuales se refiere Brousseau (1983) como obstáculos:

- Un obstáculo es un conocimiento, no una falta de conocimiento. El alumno utiliza este conocimiento para producir respuestas adaptadas a un cierto contexto que encuentra con frecuencia. Cuando se usa este conocimiento fuera de este contexto genera respuestas incorrectas. Una respuesta universal exigirá un punto de vista diferente.
- El alumno se resiste a las contradicciones que el obstáculo le produce y al establecimiento de un conocimiento mejor. Es indispensable identificar el obstáculo e incorporar su rechazo en el nuevo saber.
- Después de haber notado su inexactitud, continúa manifestándolo, de forma esporádica. Brousseau ha identificado tres tipos de obstáculos:
  - a) *Obstáculos ontogénicos ó psicogénicos*: son debidos a las características del desarrollo del niño. Por ejemplo, para comprender la idea de probabilidad se requiere el razonamiento proporcional.
  - b) *Obstáculos didácticos*: resultan de las elecciones didácticas hechas para establecer la situación de enseñanza. Por ejemplo, la introducción de un nuevo simbolismo tal como  $(\sum x_i)/n$ , cuando los estudiantes necesitan trabajar con ejemplos concretos.
  - c) *Obstáculos epistemológicos*: Relacionados intrínsecamente con el propio concepto y conteniendo parte del significado del mismo. Por ejemplo, las circularidades que se presentan en las diferentes definiciones del significado de la probabilidad (clásica, frecuencial, subjetiva) que mostraron en su día la necesidad de una definición axiomática.

Para llegar a superar dichos obstáculos es necesario que el estudiante construya la mejor concepción sobre el tema, ya que las dificultades son generadas principalmente por falencias en la comprensión de conceptos básicos necesarios. Se estudiarán ahora algunos ejemplos de lo que han reportado investigadores en torno al aprendizaje de algunos conceptos asociados a la noción de distribución.

### 3.1. Gráficas

Se pueden presentar algunos errores en el uso de representaciones gráficas y tablas de frecuencias, la lectura y comprensión de los datos es una necesidad en el área de la estadística, es así que Curcio (1989) describe tres niveles distintos de comprensión de los gráficos:

- a) “Leer los datos”: este nivel de comprensión requiere una lectura literal del gráfico; no se realiza interpretación de la información contenida en el mismo.
- b) “Leer dentro de los datos”: incluye la interpretación e integración de los datos en el gráfico; requiere la habilidad para comparar cantidades y el uso de otros conceptos y destrezas matemáticas.
- c) “Leer más allá de los datos”: requiere que el lector realice predicciones e inferencias a partir de los datos sobre informaciones que no se reflejan directamente en el gráfico.

Por ejemplo en el caso de una nube de puntos al “leer los datos” es necesaria una lectura de las escalas que se encuentran allí o hallar el valor de las coordenadas de algunos de los puntos que se encuentren allí representados. La parte acerca de “Leer dentro de los datos” podría estar referido a cuestiones de covariación, o si la relación puede ser representadas por medio de una función lineal. La predicción del valor de la coordenada  $y$ , para un valor de la coordenada  $x$  requeriría el trabajo en el nivel de “leer más allá de los datos”.

Curcio (1987) además encontró que dentro del estudio de las gráficas, deberían estar presentes los aspectos que tienen que ver con el conocimiento previo al tema que tiene que ver con el gráfico, y claro conocer el tipo de gráfico con el que se está trabajando. Además el notó que las dificultades están presentes en los ámbitos de “leer dentro de los datos” y “leer más allá de los datos”. Por otra parte reconoce la importancia de la edad en el caso de la comprensión de gráficos.

Li y Shen (1992) estudió el caso de estudiantes de secundaria, éstos muestran ejemplos acerca de errores a la hora de escoger un tipo de gráfico para determinados problemas, por

ejemplo algunos escogieron un polígono de frecuencias para representar variables, el mal uso de los software pueden llegar a generar dichos problemas, ya que se puede presentar error a la hora de escoger escalas de acuerdo con el objetivo pretendido, o al obtener los diagramas de sectores puede que éstos no estén de acuerdo a las frecuencias de las categorías, es así que los objetivos deben estar claros y así mismo correctamente enfocados al tipo de actividad que se proponga.

### **3.2. Media aritmética**

Ahora bien haciendo referencia al caso de la media aritmética, medida que posee diversas aplicaciones en varios contextos, aparentemente resulta ser simple, pero Pollastsek y cols.

(1981) describen el error que se presenta al aplicar la fórmula:  $\frac{(100 + 180)}{2}$  en el momento de resolver el siguiente problema:

*Hay 10 personas en un ascensor, 4 mujeres y 6 hombres. El peso medio de las mujeres es de 100 libras. y el de los hombres de 180. ¿Cuál es el peso medio de las 10 personas del ascensor?*

Esta es una situación en la cual se debe hallar la media ponderada, y los datos no son fácilmente identificados y puede llegar a generar en muchos casos confusión.

Otro ítem propuesto por Pollasek y cols. (1981), trata de determinar las concepciones de los alumnos universitarios sobre el valor esperado de una observación de una variable aleatoria, de la que se conoce su esperanza matemática:

*La media en fluidez verbal de una clase de un colegio es de 400. Si extraemos una muestra aleatoria de 5 estudiantes y resulta que la puntuación de los 4 primeros es de 380, 420, 600, 400. ¿Cuál sería aproximadamente la puntuación esperada para el quinto estudiante?*

La respuesta correcta a este ítem es 400, el valor esperado en la población. Sin embargo, algunos alumnos pensaban erróneamente que la puntuación del quinto sujeto debería ser tal que, sumada a las cuatro anteriores, diera una media de 400.



El análisis que se ha llevado a cabo respecto a la media está basado en como hallar dicha medida, así pues que Strauss y Bichler (1988) decidieron investigar acerca del carácter evolutivo que puede llegar a tener la noción de media en los estudiantes, teniendo en cuenta las siguientes propiedades de ésta:

- a) La media es un valor comprendido entre los extremos de la distribución.
- b) La suma de las desviaciones de los datos respecto de la media es cero.
- c) El valor medio es influenciado por los valores de cada uno de los datos.
- d) La media no tiene por qué ser igual a uno de los valores de los datos.
- e) El valor obtenido de la media puede ser una fracción (ello puede no tener sentido para la variable considerada).
- f) Hay que tener en cuenta los valores nulos en el cálculo de la media.
- g) La media es un “representante” de los datos a partir de los que ha sido calculada.

Hasta ahora se ha considerado a la media como un valor representativo dentro de un conjunto de datos, de esta manera se ha tendido a situar la media en el centro del recorrido de la distribución, sin tener en cuenta que dicha propiedad solo se va a cumplir cuando las distribuciones son simétricas, en el caso de las asimétricas van a pasar a jugar un papel mas importante la moda o la mediana, ya que la media ya va a estar en uno de los extremos.

La comprensión de la idea de “valor típico” implica, según Russel y Mokros (1991), tres tipos diferentes de capacidades:

- Dado un conjunto de datos, comprender la necesidad de emplear un valor central, y elegir el más adecuado.
- Construir un conjunto de datos que tenga un promedio dado.
- Comprender el efecto que, sobre los promedios (media, mediana o moda), tiene un cambio en todos los datos o parte de ellos.

### **3.3. Variabilidad**

Ahora se pasará a hablar de algunas características de la idea de dispersión. Ya que algunas distribuciones con la misma media o la misma mediana, pueden llegar a tener diferentes variabilidades, siempre es necesario estudiar la dispersión de los datos, y más aún cuando se están comparando dos o más muestras o poblaciones. Por ejemplo la desviación típica mide como los datos se desvían respecto de la media

Loosen y cols. (1985) hicieron un experimento, tomando 154 estudiantes de primer curso de psicología, que no habían recibido una instrucción específica sobre la dispersión, mostrándoles dos conjuntos diferentes de bloques A y B. Las longitudes de los bloques en el conjunto A fueron 10, 20, 30, 40, 50 y 60 cm. y las longitudes de los bloques en el conjunto B fueron 10, 10, 10, 60, 60 y 60 cm. Al preguntar a los sujetos cuál de los dos conjuntos presentaba mayor variabilidad, se obtuvieron las siguientes respuestas: el 50% pensó que el conjunto A era más variable, el 36% que era más variable el conjunto B y el 14% que los dos conjuntos presentaban igual variabilidad.

Los autores interpretaron estas respuestas como prueba de que el concepto intuitivo de variabilidad se equipara al de “no semejanza”, es decir, cuánto varían unos valores respecto a otros, más que cuánto varían los valores respecto a un punto fijo. En este sentido el conjunto A ciertamente debe ser considerado más variable que el B, aunque la desviación típica es mayor en el conjunto B.

### **3.4. Estadísticos de orden**

El estudio de los estadísticos de orden es de gran importancia, ya que primero que todo el análisis exploratorio de datos ha nacido a partir de ellos, y además son menos sensibles a pequeños cambios en los datos, y son la base de métodos no paramétricos que pueden ser aplicados con mayor generalidad. Este tema de estadísticos de orden presenta grandes problemas, tanto en el ámbito conceptual como en el procedimental.

Por ejemplo el cálculo de la mediana, los percentiles y el rango de percentiles se enseña diferente en el caso de variables estadísticas agrupadas en intervalos o que cuando no están agrupadas, y además estos valores van a variar dependiendo de la amplitud, algo que resulta difícil de aceptar por parte del estudiante.

Estepa (1990) observa las dificultades de los alumnos al interpretar la gráfica de frecuencias acumuladas de variables discretas, debido a que presenta discontinuidades de salto y su función inversa no es una aplicación: en esta correspondencia un punto puede tener más de una imagen, o vanos puntos pueden tener la misma imagen.

Barr (1980) llama la atención acerca de la falta de comprensión de los estudiantes sobre la mediana en un estudio llevado a cabo con estudiantes de edades entre 17 y 21 años. El 49% dio una respuesta incorrecta a la cuestión siguiente:

*La mediana del siguiente conjunto de números.*

*1, 5, 1, 6, 1, 6, 8 es*

*a) 1; b) 4; c) 5; d) 6; e) (otro valor); f) no sé:*

La mayoría de los estudiantes entienden la idea de mediana como valor central, pero no tienen claro a que secuencia numérica se refiere ese valor central. Los estudiantes pueden interpretar la mediana como el valor central de los valores de la variable, de las frecuencias o incluso de la serie de datos antes de ser ordenada.

### **3.5. Inferencia estadística**

Los estudios referidos al área de la inferencia, indican que los errores que allí se cometen tienen que ver con el empleo de heurísticas, las cuales tienen que ver con los procesos cognitivos que se utilizan para reducir la complejidad de un problema.

Cuando se intenta desarrollar, a través de la enseñanza, un razonamiento probabilístico correcto y coherente es posible encontrar que los alumnos presentan errores de concepción y de interpretación, sesgos y como dicen los psicólogos, tendencias emocionales. Las

causas de estas disfunciones pueden ser dificultades lingüísticas, falta de herramientas lógicas, dificultades para extraer la estructura matemática de las experiencias. La dificultad para comprender el concepto de azar como lo afirma Fischbein et al. (1991) se presenta principalmente con los conceptos de números racionales y razonamiento proporcional, conflicto entre las ideas de probabilidad y las experiencias de los estudiantes y su manera de ver el mundo, también un rechazo hacia la probabilidad por haberle sido presentada inicialmente en forma abstracta y formal como lo afirman Garfield y Alhgreen (1988) y por haber enfatizado en los currículos aquellos aspectos deterministas de la ciencia.

Por ejemplo, en la llamada “heurística de la representatividad” (Kahneman y cols., 1982) se estima la probabilidad de obtención de una muestra por el parecido de ésta con la población de la que proviene. En consecuencia, aparece cierta insensibilidad al tamaño de la muestra y una confianza exagerada en las pequeñas muestras, fenómeno que se conoce con el nombre de “creencia en la ley de los pequeños números”. Como ejemplo se proporciona la siguiente pregunta:

*Una cierta ciudad está atendida por dos hospitales. En el hospital más grande nacen aproximadamente 45 bebés cada día y en el hospital más pequeño nacen aproximadamente 15 bebés cada día. Como sabes, aproximadamente el 50 por ciento de todos los recién nacidos son varones, pero el porcentaje exacto varía de un día a otro. A veces puede ser mayor que el 50 por ciento, a veces más bajo. Durante un periodo de un año, cada hospital registró los días en que más del 60 por ciento de los recién nacidos fueron varones. ¿Cuál hospital crees que registró más de estos días?*

*El hospital grande.*

*El hospital pequeño.*

*Aproximadamente igual (esto es, si la diferencia entre ambos es menor del 5 por ciento).*

Muchas personas creen que la respuesta correcta debe ser la tercera, puesto que en ambos hospitales la proporción de varones es la misma (60 por ciento) y piensan que este es el único hecho de importancia para determinar la probabilidad de los sucesos requeridos. No conceden atención al tamaño de la muestra, aunque la teoría de la probabilidad nos enseña

que hay mayores fluctuaciones del valor de la proporción en las muestras pequeñas que en las muestras grandes.

Según Kahneman y cols. esta confianza excesiva en las pequeñas muestras tiene graves consecuencias en las aplicaciones de la estadística, especialmente en la investigación. El “creyente en la ley de los pequeños números” tiende a estimar a la baja la amplitud de los intervalos de confianza obtenidos, a sobrestiman la significación de sus resultados estadísticos y a esperar que los resultados obtenidos en los primeros ensayos se le confirmen en el futuro.

Otra consecuencia de la aplicación de la heurística de la representatividad sería el error denominado “falacia del jugador”. Por ejemplo, muchas personas creen que después de una racha larga de caras, es más probable obtener una cruz.

De forma general Garfield y Alhgren (1988) señalan las siguientes razones para algunas de las dificultades que surgen para el aprendizaje de la estadística:

- Algunos conceptos como el de probabilidad, correlación, necesitan del razonamiento proporcional, que ha demostrado ser un tópico difícil en diversas investigaciones.
- Existen falsas intuiciones que los alumnos llevan consigo al empezar la enseñanza. Aunque estas intuiciones son mejor conocidas para el caso de la probabilidad.
- A veces los alumnos muestran una falta de interés hacia la estadística, porque se les ha enseñado en forma muy abstracta en edades tempranas.

Hay dos razones más que posiblemente influyan en la dificultad del tema: En primer lugar, la Probabilidad y la Estadística tienen un desarrollo reciente, si se compara con otras áreas como la física o la matemática pura. Aunque en la actualidad existe una axiomática para el Cálculo de Probabilidades, comúnmente aceptada, a partir de los trabajos de Kolmogorov, no ha cesado aún la controversia sobre el significado último del término “probabilidad”, existiendo diversas escuelas: empiricistas, subjetivistas, lógicas, etc.

Esta controversia se repite en la inferencia estadística, con la polémica sobre si es posible o no el cálculo inductivo de la probabilidad de una hipótesis y si ello puede lograrse o no con la aproximación clásica o bayesiana de la inferencia (Rivadulla, 1991). Numerosas investigaciones muestran cómo las dificultades epistemológicas, que han debido ser superadas en el desarrollo histórico del conocimiento, se repiten con frecuencia en el aprendizaje del mismo.

Por otro lado, gran parte de los conceptos estadísticos han tenido su origen fuera del campo estricto de la matemática. La Estadística ha sido desde sus comienzos una ciencia interdisciplinar y las grandes etapas de su progreso han estado marcadas por aportaciones originadas a partir de la necesidad de resolver problemas en campos diversos. En la enseñanza los conceptos se presentan aislados de las aplicaciones originales. Pero cada una de estas aplicaciones aporta una parte del significado global de los mismos (Steimbring, 1990)). Así, por ejemplo, el concepto de media toma un significado diferente cuando se aplica como centro de gravedad, esperanza de vida o número índice.

En resumen, y como señala Green (1992, PG. 12): “los conceptos estadísticos proporcionan un área de exploración fascinante. Lo que parece tan obvio y sencillo a los estadísticos (términos como promedio, variabilidad, distribución, correlación, sesgo, aleatoriedad) ha sido el producto de la experiencia de varias generaciones de las mentes más capaces. Es demasiado esperar que esta herencia nos pueda ser transmitida sin esfuerzo por nuestra parte”.

En trabajo doctoral titulado la construcción del significado de la distribución normal a partir de actividades de análisis de datos (Tauber, 2003) se llega, entre otras, a las siguientes conclusiones: cuando se incorpora el ordenador en la enseñanza se introducen cambios en el significado de la distribución normal, puesto que afecta a los diferentes elementos de significado; el manejo de las tablas de la distribución y la necesidad previa de tipificación para el cálculo de probabilidades desaparecen; la gama de representaciones del

concepto se amplía notablemente, convirtiéndose en útiles dinámicos de exploración. Sin embargo también se destacan algunas dificultades: Entre ellas están:

- Interpretación de áreas en histogramas de frecuencia y problemas en el cálculo del área dentro de un intervalo, cuando ello implica el cambio de los extremos de los intervalos.
- Dificultad en discriminar los casos en que una variable cuantitativa discreta puede y no puede ser aproximada por una distribución continua y las implicaciones que esta aproximación tiene.
- Dificultad en recordar y aplicar correctamente los convenios de interpretación de los coeficientes de asimetría y curtosis.
- Dificultad en recordar y aplicar correctamente los convenios de lectura de los elementos constitutivos de un gráfico estadístico.
- Escasa diferenciación entre el modelo teórico y los datos empíricos y dificultad en distinguir cuándo el programa de cálculo se refiere a una u otra distribución, así como no discriminación entre los estadísticos y parámetros. Las actividades propuestas implican el trabajo de modelización y la discriminación de los planos empírico (datos) y teórico (modelo), que los alumnos a veces no llegan a separar y que puede explicar errores en la aplicación de la inferencia (Moses, 1992).
- Dificultad de uso de las opciones del software que pertenecen a un menú secundario y que son, sin embargo esenciales para el análisis.
- Escasa capacidad de argumentación, sobre todo de análisis y síntesis.

## CAPITULO 4

### ANÁLISIS DE LA ENSEÑANZA

#### **4.1. Introducción**

Hoy en día la enseñanza de la estadística tiene mayor demanda, gracias a que el desarrollo de la informática y la estadística han contribuido a que múltiples disciplinas como la medicina, la economía, la psicología y otras más, tengan acceso a sus diversas aplicaciones materializadas en contenidos gráficos, o poderosos sistemas de datos numéricos simbólicos. Esto ha contribuido a que en los lineamientos curriculares se propongan como conocimientos básicos el pensamiento aleatorio y los sistemas de datos.

No hay duda de la consideración de diversos elementos que se requieren para llevar a cabo un análisis de la enseñanza, entre los que se encuentran las prácticas del maestro, sus creencias y concepciones; las situaciones problemáticas que se presentan a los estudiantes y por último las cuestiones que tienen que ver con los recursos didácticos, dentro de las que se encuentran el análisis de los textos, las propuestas didácticas apoyadas en algunos casos en el uso de las nuevas tecnologías.

Dentro de esa gama de posibilidades, en este trabajo se ha escogido el análisis de textos por el hecho de que existen diversas investigaciones que asientan sobre la influencia positiva que los libros de texto ejercen sobre el rendimiento estudiantil. Por ejemplo, se pueden mencionar los trabajos realizados por Schiefelbein (1973), citado en Peña (1977), donde se concluye que si se mantienen constantes otros factores y sólo se varía la disponibilidad del texto, se observa que dicha variación influye en el rendimiento escolar, generando aumentos de 10 a 30 por ciento de los rendimientos promedios; también los trabajos del Departamento de Secretaría Educativa del Convenio Andrés Bello de la Organización de Estados Americanos, que han demostrado que a los niños que se le proporciona un libro de texto, alcanzan un mayor rendimiento que los que no lo tienen (Fienmayor, 1997); a esto se



agrega que, en un documento publicado por el Banco Mundial y titulado los Textos Escolares y el rendimiento académico, se indica que en dieciocho evaluaciones estadísticas, realizadas en diferentes países, se observó que la disponibilidad de textos era el factor que influía principalmente en la productividad (Núñez, 1998). Para Goetz y Lecompte (1988) en el análisis de textos escolares es posible establecer diferencias entre los objetivos de un programa y las formas como éstos se llevan en la práctica. También en Ortiz et. al (1996), se encuentra una posición en relación con el análisis de textos que justifica la realización de este tipo de tarea en el presente trabajo:

*Los libros de texto son un medio cuyo análisis nos proporciona un conocimiento del significado fijado dentro de una institución didáctica dada para un cierto concepto.*

*[...] un punto importante en la teorización usada es la consideración del significado de un concepto como una entidad compleja. Por ello conviene analizar separadamente sus componentes o “elementos de significado”, ya que los elementos específicos contemplados en los diversos textos podrían no ser los mismos o, incluso, alguno de ellos podría no aparecer en ninguno de los textos. Como resultado de lo anterior se generaría un sesgo en el significado presentado a los alumnos. (Ortiz et al., 1996, p. 23)*

Se ve entonces que resulta necesario analizar algunos libros de texto, para determinar, por un lado, si éstos están cumpliendo con los estándares de calidad planteados, y por otro, con el fin de acoplar ideas para proponer una unidad didáctica para la introducción del concepto de distribución estadística.

Lo que se hará será observar dos colecciones de textos. De estas colecciones, se tomarán los textos desde el nivel sexto hasta el grado once. Para el análisis se tendrá en cuenta, el esquema general que presenta el texto en todos los cursos para la presentación de los temas, y particularmente se observarán con mayor detenimiento aquellos en los cuales se encuentre presenten los conceptos y procesos asociados a las ideas de distribuciones estadísticas y de probabilidad.

Para seleccionar los textos a analizar se tuvo en cuenta, las diferentes secuencias que se da a los temas en los diferentes niveles, además los estándares están presentes en ambos textos, lo que se da gracias a la importancia que ha venido tomando el hecho de unificar los temas en la educación matemática, de acuerdo a las normas.

## **4.2. Análisis de la colección Alfa**

Cada uno de los capítulos, antes de iniciar los desarrollos temáticos, expone los logros propuestos para los estudiantes y luego presenta una reseña histórica alusiva a los temas que van a trabajar. Usualmente, se continúa con un trabajo de preparación al capítulo, en el cual se encuentran ciertas preguntas que tienen que ver con lo que se va a trabajar durante el desarrollo de los diversos temas. Estas preguntas en algunas veces están basadas en el trabajo que se ha realizado en el nivel anterior. Finalmente, en el cierre de los temas se presentan diversas aplicaciones de la estadística y la probabilidad.

Respecto a la forma en que se trabajan cada uno de los temas de esta colección, el proceso consiste en dar primero una situación que tenga que ver con el tema que se va a trabajar, muestran al menos una interpretación de dicha situación, y siguen luego con la definición formal de cada uno de los objetos que trabajan, continuando luego con una serie de ejemplos con los que se amplía el tema y se dan otras definiciones, o simplemente se profundiza sobre el tema que se está trabajando. Para terminar, usualmente se proponen una serie de problemas (sección que recibe el nombre de “aplico”) en la que lo principal es poner en práctica lo visto durante la sección. Cada dos lecciones, se encuentra un taller de competencias, en el cual se encuentran más que todo problemas acerca de los temas vistos hasta ese momento, y al finalizar el capítulo se encuentra: la evaluación de todo el capítulo, un trabajo para avanzar hacia el Icfes, un pasatiempo y un glosario. Ahora se verán algunas particularidades para cada uno de los niveles de esta colección.

### **4.2.1. Grado 6**

Este texto está conformado por 10 unidades, dentro de las cuales tres corresponden a geometría, seis a aritmética y una a estadística, el título de la unidad es ESTADISTICA y

los logros propuestos para este grado consisten en: organizar los datos utilizando tablas o diagramas de barras, interpretar el significado de frecuencia absoluta y frecuencia relativa, leer adecuadamente la información dada en diagramas estadísticos, hallar las medidas de posición en un conjunto de datos, y determinar adecuadamente la media aritmética de un conjunto de datos. Es importante señalar que dichos temas están basados en las diferentes clases de distribuciones de frecuencias que se pueden encontrar, sus diversas representaciones y algunos de los estadígrafos de descripción, los cuales son la base principal del presente trabajo.

Al inicio de la unidad se encuentra un resumen del surgimiento histórico de la estadística, nombrando en primer lugar los censos, luego las necesidades económicas y políticas de los países, continuando con una breve explicación de las dos grandes partes en las que se divide la estadística descriptiva e inferencial. Para terminar, en el texto de sexto se reseñan algunos de los principales autores que han trabajado sobre esta área. Cabe resaltar que este trabajo es de gran importancia, ya que la lectura es corta y concisa, es de gran ayuda para que los estudiantes se contextualicen y entiendan un poco la importancia y el porque del surgimiento de la estadística.

Dentro de las características de este texto se encuentran la presentación de preguntas con las cuales se pretende preparar al estudiante para dar inicio al trabajo. En particular, se incluyen cuatro preguntas. La primera pregunta tiene que ver con probabilidad, se da un problema con tres colores distribuidos en una caja en cantidades diferentes y se pregunta cual de estos tendrá mayor probabilidad de ser escogido, según los logros planteados para esta unidad la probabilidad no está presente, así pues que no sería aún concerniente este tipo de análisis. En la siguiente pregunta piden hallar un promedio, de cierta cantidad de dinero, ellos lo verán como un problema de división, pero no encontraran una relación con la estadística, pienso que éste debe ser planteado de una forma diferente. En la tercera pregunta, la idea es encontrar las posibles combinaciones que se pueden hacer con cinco objetos diferentes, en mi opinión esta pregunta también se encuentra mal ubicada ya que sería de gran ayuda, pero para dar inicio a las combinaciones y permutaciones, y en este

caso no se van a trabajar. Para terminar, en la última pregunta, se dan una serie de datos correspondientes a los pesos en kilogramos de 15 estudiantes, y se pide hallar tres cosas: el peso que mas se repite; cuantas veces se repite; y si todos pesaran lo mismo, cual sería el peso de cada uno. Las dos primeras cuestiones son útiles para dar una introducción al concepto de moda en un conjunto de datos.

Se continúa dando una breve explicación sobre las diversas aplicaciones de la estadística, y se habla de su importancia como una herramienta de investigación en todas las ciencias. Es de gran ayuda ya que todos los usos los da en el contexto en el que se esperaría que desarrolle el estudiante. Los temas tratados durante la unidad son, datos y estadísticas, frecuencias absolutas y relativas, diagramas de información estadística, frecuencias acumuladas tanto absolutas como relativas, la moda, la mediana, la media aritmética o promedio. Ahora bien, teniendo en cuenta la base del estudio de este trabajo, se empezará trabajando con la segunda sección de la unidad la cual tiene que ver con frecuencias absolutas y relativas, ya que la primera se basa principalmente en la recolección y clases de datos, que nos concierne pero en el ámbito de su uso.

El texto les presenta a los estudiantes una situación en la cuál a 35 alumnos de sexto grado se les pide llenar una tabla en la que se incluye el número de la lista de cada uno de los estudiantes, en que mes nació, y cuántos meses cumplidos tiene (ver *tabla 4.1*). Para los estudiantes sería de gran ayuda, llegar a hacer el diseño y el contenido de la tabla con datos tomados por ellos mismos, en este caso del número de la lista, el mes de nacimiento y los meses de cada uno de sus compañeros, de esta forma ellos se sentirían mejor relacionados con el trabajo que están llevando a cabo.

Siguiendo con el trabajo, se quiere saber cuántos estudiantes cumplen años cada mes, y también que parte de ellos cumplen en cada uno de los meses, así proceden a agruparlos por el mes de nacimiento, obteniendo la *tabla 4.2*. Basados en la tabla definen la frecuencia absoluta de un dato como el número de veces que se repite ese dato específico dentro de un conjunto; luego afirman que la frecuencia relativa da información sobre qué parte de la

población o de la muestra en estudio corresponde a la característica analizada, e indican que esta última se halla dividiendo la frecuencia absoluta por el número total de datos, y se puede expresar como una fracción, un decimal o un porcentaje; para terminar, definen rango, el cual es la diferencia entre el mayor y el menor valor de la característica observada en estudio.

TABLA 4.1.

TRABAJO SECCIÓN 2 ALFA 6

Número de lista	Mes de nacimiento	Meses cumplidos
1	Agosto	135
2	Marzo	132
3	Febrero	137
4	Enero	135
5	Abril	137
6	Noviembre	133
7	Octubre	134
8	Octubre	136
9	Diciembre	134
10	Septiembre	137
11	Junio	137
12	Mayo	132
13	Julio	138
14	Agosto	135
15	Septiembre	134
16	Mayo	135
17	Marzo	134
18	Febrero	135
19	Abril	136
20	Noviembre	133
21	Mayo	131
22	Julio	135
23	Agosto	133
24	Septiembre	135
25	Diciembre	136
26	Octubre	133
27	Mayo	132
28	Septiembre	138
29	Enero	136
30	Marzo	135
31	Octubre	133
32	Junio	134
33	Septiembre	135
34	Agosto	135
35	Septiembre	134

En cuanto a estas definiciones, es posible hacer una observación, es el hecho de lograr que el estudiante obtenga algún uso para las frecuencias, en este problema, por ejemplo se podría decir que si quieren dar un regalo a los estudiantes en cada uno de los meses de su cumpleaños, en cual de los meses se debería recolectar mayor cantidad de dinero, o por

ejemplo en que mes se encuentra el menor porcentaje de cumpleaños y cual es este, claro preguntas que se harían como una necesidad de construir la tabla de frecuencias.

Luego se encuentra un ejemplo basado en la observación de la *tabla 4.1.*, en el cuál se preguntan cuestiones que no necesitan ningún tipo de análisis, como son “¿Cuál es la frecuencia absoluta del mes de marzo?, ¿Cuál es la frecuencia relativa del mismo mes?, ¿Cuál es el rango?”. Además están resueltas, así que el trabajo del estudiante en este caso es completamente nulo.

TABLA 4.2

TRABAJO SECCIÓN 2 ALFA 6

Mes	No. de alumnos (frecuencia absoluta)	Parte del total de alumnos que cumple en cada mes
Enero	2	$\frac{2}{35}$
Febrero	2	$\frac{2}{35}$
Marzo	3	$\frac{3}{35}$
Abril	2	$\frac{2}{35}$
Mayo	4	$\frac{4}{35}$
Junio	2	$\frac{2}{35}$
Julio	2	$\frac{2}{35}$
Agosto	4	$\frac{4}{35}$
Septiembre	6	$\frac{6}{35}$
Octubre	4	$\frac{4}{35}$
Noviembre	2	$\frac{2}{35}$
Diciembre	2	$\frac{2}{35}$
Total	35	$\frac{35}{35} = 1$

En la parte de aplicación, se encuentran ocho problemas. Para los seis primeros se da un grupo de datos de diferentes observaciones, pidiendo en primer lugar el rango de éstos, luego organizarlos en una tabla con sus frecuencias absolutas y relativas, y para terminar se plantean ciertas preguntas, las cuales pueden ser contestadas con una simple observación de la tabla, lo cual limita al estudiante, ya que el está siguiendo un proceso mecánico, no se está incentivando el análisis.

Ya en la parte siete, se encuentra otra distribución de ciertos objetos la idea es clasificarlos en una tabla, como no están en una lista como se han presentado hasta ahora, se podría pensar que presentarían cierta dificultad en el estudiante. Y para la última cuestión se encuentra el trabajo de completar la *tabla 4.3.*, sabiendo que indica la clasificación de los 25 niños que practican diferentes deportes en un edificio

TABLA 4.3. EJERCICIO APLICACIÓN NÚMERO 8

ALFA 6

Deporte	No de niños	Porcentaje
Fútbol	13	52
Béisbol		20
Baloncesto	5	20
Natación	2	

¿Presenta algún tipo de abstracción? Para el primer caso solo es ver cuantos niños faltan para completar 25, y en el segundo dividir 2 entre 25.

La lección tres recibe el nombre de diagramas de información estadística, esta denominación en mi opinión debería ser, diagramas de frecuencias de datos, ya que lo que se representa con estos son las diferentes frecuencias y su variación en un conjunto de datos. En ésta lección, se encuentran dos diagramas de barras, uno que representa frecuencias absolutas, y otro para las relativas, los cuales son hechos tomando los datos de la *tabla 4.1.* Definen éste tipo de representación como un gráfico que muestra la variación de algunos datos en estudio de una población o de una muestra, y que puede ser vertical u horizontal, además muestran las instrucciones para su elaboración. Todo no puede ser dado a los estudiantes, podrían estar sólo las instrucciones para que ellos hagan la gráfica o viceversa.

Para el caso de los diagramas circulares, igualmente es presentado un ejemplo, en el cual se encuentra una tabla con frecuencias y sus respectivos porcentajes, continuando con las instrucciones para su ubicación en el diagrama, definiendo al final un diagrama circular como una herramienta para representar gráficamente partes o porcentajes de un total, en donde cada región es proporcional al porcentaje que aparece en la tabla. Cabría resaltar la relación de este tipo de representación con las frecuencias relativas dadas como porcentajes, lo cual se trabajó en la sección anterior.

En el trabajo de aplicación se encuentra la lectura de un diagrama de barras, la construcción de un diagrama de barras, uno circular y un pictograma a partir de una tabla en la cual se dan ciertas frecuencias absolutas, la realización de un diagrama de barras y uno circular dado un enunciado literal. Continuando es posible encontrar una tabla con dos informaciones, caso bivariado, para el cual el estudiante debe construir un diagrama de barras con el fin de comparar las variables, en este caso debería estar la especificación de que la representación debe hacerse simultáneamente. La lección cuatro trabaja con las frecuencias acumuladas, tanto absolutas como relativas, para dicho proceso, se inicia tomando como guía, la tabla dada anteriormente (4.1), se formulan un problema así:

*A un estudiante llamado Diego se le formulan dos preguntas: ¿Cuántos estudiantes del curso tienen 134 meses o menos de edad?, ¿Qué porcentaje de los estudiantes del curso tiene 136 meses o menos de edad?*

En primer lugar afirman que Diego suma el número de alumnos cuyas edades son 131, 132, 133 y 134 meses, afirmando que 16 estudiantes tienen como máximo 134 meses cumplidos, lo que inicia siendo falso, ya que son 15, presentaría problema ya que los estudiantes interesados retomarían la tabla, verificarían el error y perderían credibilidad en el libro de texto que están trabajando.

En la tabla se presentan diversos errores respecto a la original, ya que por ejemplo, un solo estudiante tiene 131 meses, 5 tienen 133 meses no 4, 6 tienen 134 no 7, en el caso de 136 meses, en la tabla original se presenta una frecuencia de 4 estudiantes y no 6, 4 tienen 137,



no 2, y 2 estudiantes tienen 138 meses, más no 1. Estos errores no son de mayor trascendencia, ya que la idea es reconocer las frecuencias acumuladas, pero se debe tener en cuenta no cometerlos, ya que podrían desviar el trabajo que se está llevando a cabo.

La definición de la frecuencia absoluta acumulada, está dada como la adición de las frecuencias absolutas anteriores con la frecuencia absoluta que le corresponde, y que esta indica el número de elementos del conjunto que tienen la característica con ese valor o uno menor. En mi opinión la definición está un poco confusa en cuanto a la redacción. Pero la explicación de ésta es clara, así que es entendible qué es y para que sirve hallar una frecuencia absoluta acumulada.

*TABLA 4.4*  
*LECCIÓN 4 ALFA 6*

Edad en meses	No de alumnos Frecuencia Absoluta	No de alumnos con cada edad o menos
131	2	2
132	3	5
133	4	9
134	7	16
135	10	26
136	6	32
137	2	34
138	1	35
Total	35	

Indican que Diego, para responder la segunda pregunta, utiliza dos procesos diferentes, que son descritos a continuación:

1. Calcula cuántos alumnos tienen 136 meses o menos, y divide esta suma entre 35, es decir  $\frac{32}{35}$ , el cociente que se obtiene quiere decir que aproximadamente el 91% de los alumnos del curso tiene 136 meses cumplidos o menos.
2. Suma la frecuencia relativa correspondiente a 136 meses de edad con las de los datos anteriores y también obtiene  $\frac{32}{35}$ .

Es de resaltar la consideración de los dos métodos para hallar la frecuencia acumulada relativa, ya que muestra las dos interpretaciones que se le puede dar a dicha noción, además mostrar la tabla (*tabla 4.5.*) con todas las frecuencias va a darle mayor visibilidad al estudiante de ésta.

TABLA 4.5. LECCIÓN 4 ALFA 6

Edad en meses	No. de alumnos (frecuencia absoluta)	Frecuencia relativa	Frecuencia acumulada absoluta	Frecuencia acumulada relativa
131	2	$\frac{2}{35}$	2	$\frac{2}{35} = 0.06$
132	3	$\frac{3}{35}$	5	$\frac{5}{35} = 0.14$
133	4	$\frac{4}{35}$	9	$\frac{9}{35} = 0.26$
134	7	$\frac{7}{35}$	16	$\frac{16}{35} = 0.46$
135	10	$\frac{10}{35}$	26	$\frac{26}{35} = 0.74$
136	6	$\frac{6}{35}$	32	$\frac{32}{35} = 0.91$
137	2	$\frac{2}{35}$	34	$\frac{34}{35} = 0.97$
138	1	$\frac{1}{35}$	35	$\frac{35}{35} = 1.00$
Total	35	$\frac{35}{35} = 1$		

En cuanto a la definición, ésta se encuentra en un recuadro así: Para cierto valor de una característica de una población o muestra, la frecuencia acumulada relativa adiciona las frecuencias relativas de los datos menores con la frecuencia relativa del dato específico. También se puede calcular mediante el cociente de la frecuencia acumulada del dato y el número total de datos. En estos dos métodos de cómo hallar la frecuencia acumulada relativa, se conjetura que el estudiante ya lo pudo haber discernido del ejemplo, por esta

razón, en mi opinión debería estar descrita más bien la finalidad de hallar este tipo de frecuencia.

Luego preguntan, como ejemplo, ¿cuál es la frecuencia acumulada relativa para 134 meses? Para lo cual basta observar la tabla, así que no hay ningún grado de abstracción por parte del estudiante en este caso.

Para terminar la lección representan mediante un diagrama de barras las frecuencias absolutas acumuladas, pero no explican nada basados en este. Para algo que podría ser útil dicho diagrama es para que los estudiantes noten ciertas regularidades, como por ejemplo que pasaría con el diagrama de barras si se variaran las frecuencias absolutas. Para la aplicación se encuentran tablas con espacios vacíos para completar con diferentes tipos de frecuencias, luego con la lectura de éstas piden construir diagramas de barras, y contestar ciertas preguntas para las cuales basta con observar la tabla, y hacer conversiones de fracciones a porcentajes o decimales.

En otras tablas, los datos dados están desorganizados y hay la necesidad de ubicarlos en el orden correcto de menor a mayor para realizar el trabajo con las preguntas que continúan a partir de estos sobre frecuencias acumuladas. Esto será un trabajo interesante, ya que el estudiante se fijará en este detalle, el cual resulta ser de gran importancia para el trabajo. Además se notará para que puede ser útil el trabajo con dichas frecuencias. En el punto 10 se encuentra un trabajo ya un poco diferente, se induce al estudiante a construir una muestra de datos, y hallar la frecuencia de una característica de éstos, y a partir de esta hallar las frecuencias relativas, relativas acumuladas y absolutas acumuladas, y para el final realizar el diagrama de barras para dichas frecuencias. Hacen varios diagramas de barras, pero ¿Cuál es el fin de éstos? Sería útil la inclusión de preguntas que logren que el estudiante saque conclusiones a partir de éstas. Antes de la siguiente lección es posible encontrar un Taller, en el cuál se encuentran tres fases de trabajo, una individual, otra en grupo y para terminar una de profundización, el sistema de trabajo de éstas es similar al de los talleres, así que estaría demás nombrarlos o explicarlos. Finalmente, las lecciones 5, 6, y 7 tratan

algunos de los estadígrafos de descripción, los cuales no juegan un papel de mayor importancia en el desarrollo del presente trabajo.

#### **4.2.2. Grado 7**

El título de la unidad es ESTADÍSTICA Y PROBABILIDAD y los logros propuestos para este grado están divididos por procesos de la siguiente manera:

Conexiones: Relacionar diferentes representaciones que permiten reconocer situaciones estadísticas y aleatorias en eventos cotidianos; reconocer la presencia de las matemáticas en contextos no matemáticos y las aplicaciones para predecir resultados de eventos repetitivos y de resultados al azar, con base en eventos repetitivos y aleatorios; organizar en tablas, diagramas y gráficas los datos presentes en una situación real;

Comunicación: Relacionar diagramas y gráficas estadísticas con las medidas estadísticas y las nociones aleatorias; interpretar y presentar en forma oral y escrita el significado de la información de tablas y diagramas de barras y lineales, que permiten establecer el comportamiento de un suceso; presentar y dar sentido mediante diagramas y gráficas a las situaciones estadísticas que se pueden detectar con los diferentes contextos.

Razonamiento Lógico: Comprender cómo se da el razonamiento estadístico y aleatorio, como utilizarlo en la comprensión y solución de situaciones matemáticas y no matemáticas; justificar los procedimientos para calcular medidas estadísticas, y da razones del porque una cierta medida; formular conjeturas y desarrollar procedimientos para generalizarlas y validarlas.

Resolución de Problemas: Resolver problemas de diferentes contextos, adaptando procedimientos estadísticos y aleatorios; combinar diversas estrategias y procedimientos en la resolución de problemas; predecir el resultado de un problema, mediante el uso de propiedades de las medidas estadísticas y de los diferentes tipos de probabilidad; aplicar algoritmos en la determinación de medidas estadísticas, en problemas de tipo estadístico.

Así pues a diferencia del texto del nivel anterior en este caso es más explícito el trabajo que se va a llevar a cabo, continuaremos viendo, si será posible que los estudiantes lleguen a lograr lo planteado con el desarrollo de la unidad.

El texto sigue con una breve reseña histórica que se refiere a los estudios de mortalidad y natalidad hechos por Edmund Halley, continua con Cantor y la creación de una nueva disciplina de la teoría de conjuntos hacia finales del siglo XIX, y termina con el surgimiento de la teoría de las probabilidades con la posible solución de un juego de dados y sus apuestas. Como vemos, a diferencia del texto anterior, en este caso se empieza a tocar el tema del surgimiento de la probabilidad, y además se da una continuidad en cuanto a tiempo y espacio en la historia. Las dos primeras preguntas de preparación se basan en conjuntos. Cabe anotar que es importante que la misma esté ubicada en este lugar ya que para trabajar durante el capítulo será necesario utilizar este conocimiento adquirido durante los cursos anteriores. En la última pregunta se encuentra una tabla, como la que se ilustra a continuación (*tabla 4.6.*):

TABLA 4.6.

ALFA 7

No de alumnos ausentes	1	2	3	4	5
No de días en que se dio ese número de ausencias	6	9	3	2	1

Basados en esa tabla se plantean las siguientes preguntas:

- ¿Cuál fue el mínimo número de ausencias?
- Si todos los días hubiera faltado el mismo número de alumnos, ¿Cómo sería la tabla?
- ¿Cuál es el número de fallas que se presentó más frecuentemente en los días del mes anterior?
- En un calendario escribe el número de alumnos ausentes por día, usando la información de la tabla.

En cuanto a la primera cuestión, se considera que esta pregunta es para determinar si los estudiantes diferencian entre la frecuencia y la variable que se está trabajando, ya que al referirse a número en ambos casos, y en la misma pregunta, los estudiantes deben hacer un análisis de cual de los dos indica el mínimo y el máximo número de ausencias, puede que ellos contesten que fue 1 y 5, o que al contrario se refieran al número de días, caso en el cuál podrían referirse a que el mínimo número de ausencias fue 5 y el máximo 2.

En el segundo caso lo que podrían hacer los estudiantes, sería sumar el número de fallas, y dividirlo entre 5, y este número ubicarlo en cada uno de los correspondientes días. Otro caso que podría presentarse, es que ellos sumaran el número de días y lo dividieran por 5 igual que el número de alumnos, pero el número que les da en el caso de los días no es divisible entre 5, de esta forma se encontraría una dificultad en este tipo de problema.

Para la tercera parte, en mi opinión, el hecho de referirse “al mes anterior” puede producir cierta confusión en el estudiante, ya que tal vez él al leer el problema no le preste atención a esta cuestión, pero no hay duda que se está indagando sobre la moda, que en este caso será 2. Lo cual podrán responder teniendo las bases del año anterior. El fin de esta pregunta no es claro, pero seguramente tendrá muchas respuestas al respecto, ya que la frecuencia puede ser ubicada como cada uno de los estudiantes así lo crea conveniente.

De manera general se podría afirmar que la formulación de las preguntas presentaría confusiones en los estudiantes, entre las variables y las frecuencias, desviando de esta forma el fin principal de cada una de las cuestiones que se tocan en las preguntas, que son para guiar el trabajo de obtener la media, la moda y la mediana.

En cuanto a las aplicaciones, ya se refieren a la estadística combinada con la probabilidad, afirmando que estas son una buena herramienta para estudiar y predecir comportamientos de poblaciones, las cuales son estudiadas por áreas de las ciencias naturales como son la biología, física y química y por las ciencias sociales como la psicología y la sociología; haciendo énfasis también en la era tecnológica actual, gracias a la cual es posible utilizar

diversos programas con los cuales se pueden calcular medidas estadísticas y probabilidades, y además predecir la forma de actuar de las variables poblacionales en un futuro. Incentivando de esta forma a los estudiantes que se inclinan por distintas áreas al uso de la estadística y la probabilidad.

Los temas que se tratan en cada una de las lecciones son: clases de variables y tablas de frecuencias; media aritmética, intervalo modal e intervalo mediano; subconjuntos o partes de un conjunto; permutaciones y combinaciones de los elementos de un conjunto; fenómenos aleatorios; introducción a la probabilidad. Las clases de variables son de gran importancia. Deberían ser abordadas antes de iniciar con el estudio de distribuciones de frecuencias, son la base del trabajo que se lleva a cabo, ya que sobre estas es que se realiza cualquier análisis de frecuencias, en mi concepto este tema debería estar en grado 6° antes de iniciar con cualquier trabajo propio de la estadística.

En este primer tema, lo que se hace es dar dos ejemplos con los cuales se intenta dar una explicación de lo que es una variable cualitativa y una variable cuantitativa, y como ésta última puede ser de dos clases continua o discreta, continuando con su definición, dada de tal forma que para un estudiante de 7° grado resulte fácil de entender.

Luego se habla de las representaciones de dichas variables, diciendo que si se trata de una variable cualitativa se utilizan tablas, diagramas de barras separadas, o diagramas circulares, y si es una variable cuantitativa discreta esta se representa usando tablas de frecuencias, diagramas de barras separadas y diagramas circulares; sin hacer énfasis en que lo que se está mostrando son las frecuencias de las diversas características, además que se pueden utilizar otros tipos de gráficas que ellos no han nombrado hasta el momento.

Después de estas aclaraciones continúan con un ejemplo, en el cuál ya se encuentra presente la palabra frecuencia y frecuencia relativa, y su respectivo diagrama circular, ya este tema se había trabajado en el grado anterior ¿Por qué repetir? Tal vez porque en este

caso ya se está haciendo énfasis en las clases de variables, he aquí la importancia de haber trabajado desde un inicio con las clases de frecuencias.

La representación de una característica cuantitativa continua, se hace mediante esta agrupada en clases o intervalos numéricos, por medio de diagramas de barras unidas, llamadas histogramas, y usando diagramas lineales. En el texto anterior, se trabajó con histogramas, pero nunca se hizo una aclaración como la de ahora, acerca de su diferencia con el diagrama de barras. Continuando con un ejemplo acerca de estaturas tomadas a ciertos alumnos. En una primera presentación el maestro agrupó los datos como se dan en la *tabla 4.7*.

TABLA 4.7.

LECCIÓN 1 ALFA 7

Estatura(cm.)	146	148	150	152	153	155	157	161	163	166	168	170
No de alumnos	1	2	2	3	5	9	7	4	3	2	1	1

Aquí ya se tiene una forma de clasificar los datos a la cual no se hace mención en la explicación anterior a ésta, entonces va a presentar confusión en los estudiantes, ya que no se ha hablado acerca de ésta. Ahora dicen que como la estatura es una característica continua, es posible agrupar los datos, considerar intervalos de cinco centímetros y reunir, en cada uno, los alumnos cuya estatura está en ese intervalo, así (*tabla 4.8*):

TABLA 4.8. LECCIÓN 1 ALFA 7

Estatura Intervalos de estatura	No de Alumnos Frecuencia Absoluta
146 – 150	5
151 – 155	17
156 – 160	7
161 – 165	7
166 – 170	4



A continuación se presenta una explicación de dicha distribución de los intervalos, como el rango de estatura está entre 146 cm. y 170 cm., podemos agrupar en 5 intervalos de 5 cm. cada uno. El maestro agrupa así: en el primer intervalo incluye a los dos estudiantes con 145.5 cm. o más hasta 150.4 cm.; en el segundo a los que tienen entre 150.5 y 155.4; etc. En este punto el estudiante seguramente se preguntará el porque de estos intervalos, pero más adelante se encuentra la explicación, pasemos a analizarla.

*Para conocer la longitud de un intervalo se encuentra la diferencia entre los dos valores extremos inferiores o superiores de dos intervalos consecutivos. La representación gráfica llamada histograma, se realiza mediante barras unidas, con base proporcional a la longitud del intervalo y altura proporcional a la frecuencia del intervalo. El diagrama lineal se obtiene uniendo con segmentos los puntos medios de las bases superiores de las barras del histograma.*

Se encuentran diversos problemas en la anterior explicación, uno de ellos es que la longitud de cada uno de los intervalos se debe conocer antes de establecer éstos, la idea es ayudarles a los estudiantes a encontrar un criterio para formar los intervalos dependiendo ya sea de la cantidad de los datos, o de su orden. También falta aclarar, lo que es la marca de clase, y como hallar ésta, la cual juega un papel importante en la construcción del histograma y del diagrama lineal, como ellos lo llaman, el cual se denomina en la mayoría de los libros de texto de estadística, polígono de frecuencias. La denominación resulta en cierto grado importante, ya que a la hora de leer información en un periódico o en otro libro de texto es necesario que los estudiantes estén contextualizados, tal vez ellos relacionan el diagrama lineal con el polígono de frecuencias, pero podría causar confusión en algún momento.

Para finalizar el trabajo de la sección, antes de la aplicación, se muestra un último ejemplo, en el cual se encuentra una tabla con los pesos de 40 alumnos, y se hace la aclaración de hacer una nueva tabla usando intervalos de 7 Kg. cada uno, y se pide además construir el histograma y el diagrama lineal correspondiente. En este caso insisto en la necesidad de incentivar al estudiante una forma para que el mismo halle el ancho de los intervalos, dependiendo de la cantidad de datos, hay procedimientos para hacerlo, estos fueron presentados en el anterior capítulo. En la aplicación, lo que se hace en general es mostrar

tablas y con base en estas se plantean ciertas preguntas, de tipos de variables, rango de variación, se pide representar los datos ya sea por medio de histograma o diagrama lineal, o agruparlos en intervalos dados. También se tienen en cuenta la frecuencia relativa, diagrama circular temas que fueron vistos durante el curso anterior, claro si se sigue con el mismo libro de texto, y en este nivel son usados en uno de los ejemplos. Algo de resaltar es que piden analizar las gráficas, planteando algunas preguntas sobre los datos, las cuales son posibles de deducir con la ayuda de las gráficas.

#### **4.2.3. Grado 8**

El título de la unidad es ESTADÍSTICA Y PROBABILIDAD y los logros propuestos para este grado consisten en: calcular e interpretar el significado de las medidas de tendencia central, media, mediana y moda; determinar e interpretar algunas medidas de dispersión de un conjunto de datos; interpretar y formular inferencias; argumentar con base en las medidas de tendencia central y de dispersión de un conjunto de datos; y establecer la probabilidad de ocurrencia tanto de eventos simples como de eventos compuestos.

En cuanto al surgimiento de la estadística, lo que hacen en este texto es continuar con la historia dada en el grado anterior, ya en este caso dan inicios de la enseñanza universitaria, y la necesidad de la estadística en la sociedad a finales del siglo XVII y comienzos del siglo XIX, además se hace énfasis en la estadística moderna. También se habla de Mendel y Galton quienes aplicaron métodos estadísticos y probabilísticas al estudio de la herencia y la desarrollaron modelos para la regresión y correlación estadísticas, para terminar se encuentra Pearson quien inició y aplicó muchos métodos estadísticos que hoy se usan, como lo es la distribución Chi Cuadrado. Como se puede notar en este apartado de la historia se tratan temas un poco más avanzados tanto de la estadística como de la probabilidad, los cuales al ser leído por los estudiantes puede presentar algunos interrogantes por parte de los estudiantes, los cuales podrán ser resueltos, ya sea durante el desarrollo de la unidad, o en el curso posterior.

En las preguntas de preparación se trabaja con promedios, definición de media aritmética, mediana y moda, las cuales podrán ser resueltas por aquellos estudiantes que han seguido este mismo libro de texto durante los grados anteriores. Además se indaga sobre las diferencias entre variable cualitativa y cuantitativa, variable discreta y variable continua; y acerca de aleatoriedad y probabilidad básica, temas que fueron trabajados durante el desarrollo del trabajo en estadística del grado séptimo. Las aplicaciones de las que hablan, diferentes a las de los grados anteriores son los censos y su importancia en la política, además acerca de los procesos biológicos y químicos de la física que se rigen por reglas estadísticas que establecen las expectativas de sus resultados. Y para terminar como la estadística y la probabilidad se usan en los juegos de azar.

#### **4.2.4. Grado 10**

El título de la unidad es PROBABILIDAD, y los logros propuestos para este grado son:

**Conexiones:** A partir de los principios básicos del conteo determina el número de casos que cumplen una o varias condiciones para la conformación de un conjunto de interés, utiliza la función de probabilidad para resolver problemas de distintas áreas con organización de comités, salud, control de calidad y juegos de azar.

**Comunicación:** Explica a sus compañeros el espacio muestral asociado a un experimento aleatorio, da argumentos acertados respecto a la probabilidad asignada a un evento donde el número de casos favorables corresponde a una combinación o a una permutación de casos, explica satisfactoriamente la condición de dependencia e independencia entre dos eventos, y construye un árbol de probabilidad para hallar probabilidades condicionales.

**Razonamiento lógico:** Da las justificaciones correspondientes de acuerdo con los axiomas de probabilidad y resultados establecidos en la deducción de nuevas propiedades de la función de probabilidad, justifica utilizando los principios de conteo establecidos, razonamientos fundamentados en el esquema de las casillas, y justifica utilizando las leyes establecidas de probabilidad argumentos de probabilidades condicionales fundamentadas en árboles de probabilidad.

Resolución de problemas: Aplica correctamente los principios fundamentales del conteo en situaciones que implican selección de muestras, determina los casos posibles y los casos favorables de un evento aleatorio para hallar la probabilidad del evento, y utiliza las propiedades de la función de probabilidad para determinar la probabilidad de un evento.

En la parte histórica se hace referencia al desarrollo de la teoría de probabilidades la cual es atribuida principalmente a los jugadores y matemáticos europeos del S. XVII. Meré jugador francés con la ayuda de Pascal (1623 – 1662) obtuvo la probabilidad de un juego de azar. En general el cálculo de la probabilidad nació en 1654 con Pascal y Fermat al encarar juegos de azar, uno de ellos consistía en averiguar el número de veces que debían lanzar dos dados para que la probabilidad de obtener un número seleccionados fuera el 50%, el resultado fue hallado por métodos que hoy conocemos como análisis combinatorio, ellos no publicaron sus trabajos, quien lo hizo fue Huygens (1629 – 1695) junto con los suyos. Es extraño que la probabilidad haya surgido en una época tan reciente, porque es indudable que la humanidad siempre ha estado interesada en evaluar la incertidumbre.

En el caso de los problemas de preparación, primero se da un ejercicio de permutaciones y combinaciones, pero la idea es que el estudiante indague y trabaje experimentando sin fórmulas. En la segunda parte hay un ejercicio de unión e intersección de conjuntos, lo cual es aplicado en la probabilidad, es indudable que los estudiantes no le verán la aplicación a la estadística, pero les será de gran ayuda para el trabajo que se lleva luego en la unidad. Para la aplicación se nombran los casos de el análisis que debe hacer un actuario de la tasa de reclamaciones por siniestros en el ramo de los seguros para automóvil, las predicciones que hace un epidemiólogo sobre la frecuencia relativa del número de pacientes que responden positivamente a un tratamiento para determinada enfermedad, los estudios que realiza un ingeniero de telecomunicaciones respecto al flujo de información por unidad de tiempo que se dará en cierto canal de transmisión, los vaticinios que plantea un economista respecto al mercado de divisas, etc., y presentan una tabla de esperanza de vida la cual se encuentra explicada.

Es posible ver que las aplicaciones en este texto varían, ya se están refiriendo a temas más avanzados en el área de la estadística y la probabilidad, además la tabla de esperanza de vida, está en contexto con los estudiantes.

Los temas que se tratan en las lecciones de esta unidad son, los espacios muestrales, los principios fundamentales del conteo, el concepto de probabilidad y probabilidad condicional. De esta forma el tema que de alguna nos interesa es de los espacios muestrales, ya que este será relevante en el desarrollo de las distribuciones de probabilidad.

Lo que hacen en primer lugar para la lección 1 de espacios muestrales es dar la definición de los modelos determinísticos y los probabilísticos, y se presentan ejemplos de cada uno de ellos, luego definen experimento aleatorio, y espacio muestral, de una manera que sin duda los estudiantes entenderán, continuando con ejemplos en donde se manejen estos dos términos, como se puede ver se utiliza el método tradicional. Para terminar se trabaja con eventos simples y compuestos, con los cuales se trabajan diferentes propiedades de unión e intersección asociadas a la probabilidad.

En el trabajo de aplicación se trabaja con conjuntos, tema que tal vez ya ha sido visto por los estudiantes; si no ha sido así sería necesario hacer una retroalimentación de éste durante el trabajo de preparación, también se pide clasificar diversos experimentos en probabilísticos y determinísticos, describir espacios muestrales de experimentos, hasta este punto el estudiante no ha tenido que hacer ninguna clase de análisis ya que con la lectura de los ejemplos y las definiciones puede que los resuelva.

#### 4.2.5. Grado 11

El título de la unidad es **MODELOS PROBABILÍSTICOS** y los logros propuestos son:

Conexiones: Resolver problemas de distintas áreas en donde sean requeridos los modelos probabilísticos, encontrar valores de costos esperados, tiempo esperado, número esperado de sobrevivientes etc., donde ha definido claramente la variable aleatoria y el modelo

utilizado y determina la probabilidad de algún suceso medido por una variable aleatoria continua e interpreta los resultados obtenidos.

Comunicación: Diferencia la noción entre variable aleatoria discreta y continua, dando explicación de sus características, da argumentos formales del modelo probabilístico discreto que debe utilizarse en la solución de algún problema, explica satisfactoriamente las conclusiones alcanzadas luego de aplicar algún modelo probabilístico en la solución de un problema y justifica las diferencias alcanzadas al aplicar el modelo binomial y de Poisson.

Razonamiento lógico: Realiza deducciones para hallar los valores esperados y varianzas de algunas variables discretas y continuas, utiliza pasos lógicos para hallar la probabilidad de un evento en modelos discretos, al emplear la tabla de distribución de probabilidad, y explica las propiedades generales de una función de distribución de probabilidad o de las particularidades de algún modelo.

Resolución de problemas: Resuelve problemas con variables aleatorias discretas atendiendo al valor esperado y la probabilidad de algún evento, resuelve problemas de modelos discretos, realizando cálculos numéricos con la calculadora o tabla de probabilidades en el apéndice y resuelve problemas de modelos continuos, apoyándose en el cálculo integral y en la tabla de valores de la distribución normal estándar.

En la parte de la historia lo que se hace es hablar de cómo el azar surgió hasta el siglo XVI porque la teoría incipiente de las matemáticas no permitía que este se desarrollara antes, quien dio inicio fue Cardano (1501 – 1576) a quien se debe el primer escrito de probabilidad referido al juego de los dados y Demoivre (1667 – 1754) muestra el teorema del límite central al que Gauss y Laplace dieron luego una forma más general.

Para el caso del trabajo de preparación lo que se hace es trabajar con experimentos aleatorios en los cuales se pueda observar alguna regularidad y así llegar a una conclusión, con este trabajo el autor podría adentrarse al área de la inferencia estadística, pero en este

caso solo le pide al estudiante que halle el espacio muestral lo que no conlleva mayor análisis, sino a utilizar lo que trabajó en el nivel anterior. En el siguiente apartado se dan ciertas probabilidades en porcentajes, de ganar ciertas cantidades de dinero, y se pide al estudiante que halle un cierto valor de ganancia dependiendo de los valores dados, lo cual será posible de resolver, haciendo un análisis y resolviendo un par de divisiones y multiplicaciones.

Empezando con el contenido, en este tema se puede ver que el trabajo con los conceptos y los diferentes ejemplos son extensos, tal vez por la complejidad de éstos. Las temáticas tratadas son variables aleatorias, modelo binomial, modelo de Poisson, Modelo normal. En la primera sección titulada variables aleatorias, se inicia dando una definición de lo que es una variable aleatoria, la cual está así *“Es una variable que toma diferentes valores numéricos que han sido determinados por el azar”*. La definición en mi concepto es correcta, pero faltaría aclarar que estos valores han sido dados por el suceso aleatorio, lo de la función resultaría bastante complejo para los estudiantes en este caso. Se sigue con la definición de variable aleatorias (continua y discreta),. Veamos como lo hacen:

Como se verá en la *gráfica 4.1*. la idea principal es hallar la probabilidad de diversos valores, definiendo la distribución de probabilidad como un lista, fórmula o gráfica que asigna una probabilidad a cada valor de la variable aleatoria.

Se continúa con la definición de la distribución de probabilidad acumulativa para variables discretas, y con su respectivo ejemplo, en el cual se da una tabla para la cual se debe verificar si cumple las propiedades para ser una distribución de probabilidad, además se pide hallar la distribución de probabilidad acumulativa, hallar otras dos probabilidades, y trazar una grafica, para la cual el estudiante hasta ahora no tiene idea, no hay duda que este ejemplo debe ser explicado con sumo cuidado al estudiante, ya que entran conceptos que no son muy asequibles, y que se están trabajando por primera vez.

GRÁFICA 4.1.

ALFA 11

Ejemplo 5

Lanzamos una moneda tres veces en forma consecutiva. Analicemos las probabilidades asociadas a los eventos de este experimento.

Espacio muestral:  $EM = \{(C, C, C), (C, C, S), (C, S, C), (S, C, C), (S, C, S), (S, S, C), (C, S, S), (S, S, S)\}$ .

Probabilidad definida en EM: si suponemos que en cada lanzamiento los resultados esperados —cara (C) o sello (S)— son igualmente posibles, entonces la probabilidad de que aparezca C es  $\frac{1}{2}$  y de que obtenga S también es  $\frac{1}{2}$ , luego la probabilidad de cada uno de los ocho eventos simples del EM es  $\left(\frac{1}{2}\right) \times \left(\frac{1}{2}\right) \times \left(\frac{1}{2}\right) = \frac{1}{8}$ .

Observemos que el número de caras que aparece en cada lanzamiento no puede determinarse, pues varía de manera aleatoria; por tanto, si denotamos con  $X$  la variable “número de caras que aparecen”, entonces los valores posibles que puede tomar esa variable son 0, 1, 2 o 3. Luego tenemos las siguientes implicaciones:

Si  $A_0 = \{(S, S, S)\}$ , entonces  $P(A_0) = P(X = 0) = \frac{1}{8}$ .

Si  $A_1 = \{(S, C, S), (S, S, C), (C, S, S)\}$ , entonces  $P(A_1) = P(X = 1) = \frac{3}{8}$ .

Si  $A_2 = \{(C, C, S), (C, S, C), (S, C, C)\}$ , entonces  $P(A_2) = P(X = 2) = \frac{3}{8}$ .

Si  $A_3 = \{(C, C, C)\}$ , entonces  $P(A_3) = P(X = 3) = \frac{1}{8}$ .

donde  $P(A)$  debemos leerlo la probabilidad del evento  $A$  y  $P(X = k)$  la probabilidad de que la variable aleatoria  $X$  asuma el valor  $k$ . ▲

Para una variable aleatoria discreta podemos listar las probabilidades de los distintos valores que esta asume; eso podemos lograrlo mediante una tabla de valores, una fórmula o una gráfica, que asigne una probabilidad a cada valor de la variable aleatoria.

Luego se define lo que es el valor esperado, la varianza y la desviación estándar con base en fórmulas, en las cuales hay letras que el estudiante no ha visto hasta ahora, continuando con un ejemplo en el cual se hallan dichos estadígrafos de descripción. En el último ejemplo de la sección se da una función de probabilidad, en el cual se piden hallar los estadígrafos, pero aún el estudiante no ha estudiado la función como tal. De esta forma se puede ver que el trabajo a realizar paralelo al texto es extenso, ya que en éste se avanza sin tener en cuenta que todos esos conceptos son bastante complejos, sobretodo para el estudiante, que ha venido trabajando durante los anteriores cursos con estadística y probabilidad básica.



En la siguiente sección se trabaja con el modelo binomial, la explicación e introducción a éste, se encuentra muy bien explicada ya que da los parámetros para que una distribución sea de Bernoulli, y continuando se da un ejemplo en el cual se da el suceso, el éxito y por consiguiente el fracaso, en donde se pide encontrar la distribución de probabilidad, en la cual se obtiene una tabla, la cual no es difícil de entender. Ahora bien se refieren a que el modelo binomial se basa en un experimento en el cual se lleva a cabo una sucesión finita de ensayos de Bernoulli, y luego lo definen dando las condiciones que debe cumplir un experimento para ser binomial.

En el siguiente ejemplo se presentan cinco experimentos en los cuales se pide hallar si estos son de tipo binomial, y continúan con la deducción de la distribución de probabilidad de una variable aleatoria binomial como se hace en la mayoría de libros de texto, la que se encuentra en el capítulo 1 del presente escrito, con su respectivo ejemplo. No hay duda que en este espacio debería haber más libertad para que el estudiante construya.

En la lección tres titulada modelo de Poisson, se tiene en primera instancia la fórmula dada para ésta distribución, la cual para un estudiante no será muy clara, ya que él se preguntará sobre su deducción, y de esta manera se presentan tres ejemplos en los cuales la aplican. En la cuarta lección se encuentra el modelo normal, el cual se trabaja en contexto continuo y discreto, y a partir de ciertos ejemplos definen lo que es una función de densidad, para la cual dan sus características, más no lo que esta es.

En general el trabajo que se hace con distribuciones probabilísticas es bastante ambicioso, ya que sería más conveniente, desde mi perspectiva, trabajar con solo dos de ellas, y de ésta forma intentar que queden claras para los estudiantes.

### 4.3. Análisis de la colección Prentice Hall

Cada capítulo, antes de iniciar las temáticas, da los logros propuestos y los correspondientes indicadores de logro, luego expone una conexión con la historia. Comparando con la colección anterior, en este caso esta colección se remite a una época

más antigua y más concisa, pero es muy diferente de un texto al otro, más adelante veremos por que. Se continua con los requisitos, aquellas preguntas son necesarias para que el estudiante se enfrente al trabajo del capítulo en este caso no se presentan preguntas explícitas respecto al capítulo que se va a trabajar, más bien se intenta retomar aspectos más bien del área de matemáticas los cuales son mas asequibles para el estudiante, digamos que en este caso el estudiante está mas limitado a pensar.

La generalidad de este texto en la forma en que trabajan cada uno de los temas, consiste en dar en primer lugar un ejemplo que tiene que ver con el tema que se va a trabajar, y muestran su solución, siguiendo con una o más conclusiones las cuales ayudan al estudiante a darse una visión más amplia y explicativa acerca de lo que se esta trabajando, en este texto a comparación del texto anterior se hace menos extenso éste trabajo, y las explicaciones la mayoría de veces son entendibles, porque se manejan con un lenguaje claro y conocido para el estudiante; luego se encuentra el trabajo de practica, como su nombre lo dice el estudiante va a utilizar lo que entendió en la sección anterior, en esta se encuentran máximo cuatro situaciones para realizar individualmente, y una para llevar a cabo en grupo, esta última es así porque el trabajo resulta ser un poco más largo.

Para terminar el capitulo se encuentra una actividad de nivelación en la cual se trabaja con todos los temas vistos durante éste, luego se pueden ver las situaciones problema, en las cuales se trabajan ejercicios muy parecidos a los anteriores, más adelante en un examen más detallado se verá que la diferencia no es mucha, además se tiene un taller titulado El Nuevo Examen de Estado, en el cual como su nombre lo dice, se encuentran preguntas que pueden ser encontradas en dicho examen basadas en la estadística, siguiendo con la Gimnasia matemática que consiste en que el estudiante dé solución a situaciones problema, en este caso si se pueden encontrar para algunos niveles, formar rompecabezas y podrá resolver sopas de letras, para terminar se encuentra una sección titulada tecnología aplicada, en la cual guían al estudiante para utilizar el programa derive y hallar diversas medidas estadísticas, trabajo que resulta ser muy interesante tanto para el maestro como para el estudiante, se tendrán otras visiones acerca de ésta área.

### 4.3.1. Grado 6

Los logros e indicadores de logro correspondientes para este grado en estadística son:

#### **En conocimiento matemático:**

- Comprende y usa la media, la mediana y la moda en un conjunto pequeño de datos y saca conclusiones estadísticas.
- Identifica las medidas de tendencia central; diferencia moda, media y media.
- Determina a partir de un conjunto de datos la frecuencia absoluta, la frecuencia acumulada, y la frecuencia relativa.
- Establece diferencias entre frecuencias absolutas y frecuencias acumuladas.
- Determina las analogías entre porcentaje y frecuencia relativa.

#### **En situaciones - problemas:**

- Interpreta y usa datos estadísticos en diferentes contextos.
- Analiza gráficas estadísticas y plantea situaciones que puedan expresarse por medio de tablas o gráficas estadísticas.

#### **En comunicación matemática:**

- Expresa las situaciones planteadas, en las diferentes formas de representación usadas en las matemáticas.
- Presenta la información en tablas y en diagramas de barras, circulares, de líneas y pictogramas.

De esta forma se puede notar que los logros e indicadores son los esperados para llegar a un desarrollo óptimo por parte del estudiante del tema que se está trabajando, pero se verá que en el desarrollo del trabajo no se incentiva a éste para algunos de los logros. Por ejemplo el estudiante va a usar y a hallar la media, pero no la va a llegar a comprender, en cuanto al logro de las situaciones problema, el contexto que se usa, será necesario un contexto adecuado para el estudiante.

La conexión con la historia se refiere a de donde viene la palabra estadística, además se refiere a como en la época de la Biblia, hubo censos realizados por Moisés y por David.

Además alrededor del año 3000 a.C. se cuenta que un rey llamado Yao hizo levantar un recuento agrícola, y para terminar acerca de cómo durante el imperio romano se hicieron los denominados Renta Vitalicia, con el fin de calcular la vida media a distintas edades. Se puede notar, comparándolo con el texto Alfa que en este caso se remiten un poco más al pasado.

En el caso de los requisitos se pueden encontrar cuatro preguntas para la primera de ellas es necesaria una reflexión acerca de para lo que podría servir la estadística y el porque de su importancia, en segundo lugar hay un trabajo más de campo en el cual el estudiante tendrá la oportunidad de interactuar con sus compañeros por medio de una encuesta en su colegio, con estos datos que recojan tendrán que hacer la gráfica de barras más apropiada para éstos, y para terminar hacer un análisis escrito de lo que pueden encontrar después de la encuesta y la gráfica. Este trabajo en mi concepto es interesante, ya que el estudiante estará más contextualizado en lo que es y para lo que le puede servir la estadística en determinado momento, logrando así llamar su atención a la hora de abordar los diversos temas.

Los temas que se trabajan son la estadística, la frecuencia absoluta, las graficas de barras, los diagramas circulares, los pictogramas, la moda y la mediana, la media, la frecuencia relativa y la frecuencia acumulada. En la primera parte se hace una breve introducción de lo que es la estadística descriptiva y la estadística inferencial, la explicación es bastante breve, y en cuanto a lo que se refieren de estadística inferencial, no hablan acerca de lo que es la probabilidad, lo que hacen es aplicar una regla de tres, lo que no es la probabilidad.

Se continúa con lo que es una frecuencia absoluta mostrando una tabla con las temperaturas de ciertos días, y a partir de ésta se halla una tabla de frecuencias absolutas, y luego se define lo que es ésta. En la práctica se dan ciertos grupos de datos con el fin de que los estudiantes hallen las respectivas frecuencias absolutas, también se muestra una tabla de frecuencias y se plantean preguntas sobre ésta, luego se encuentra una tabla la cual expide completar, en el caso de los días esta va ascendiendo de uno en uno, pero para el caso de la frecuencia no hay un patrón a seguir, así pues que se podría tener cualquier valor.

El trabajo en grupo consiste en realizar una encuesta dentro del mismo, y organizar los datos en una tabla de frecuencias, actividad que es más dinámica y productiva para los estudiantes.

Continuando con el texto se encuentran las graficas de barras, para las cuales se presenta un ejemplo con una tabla de frecuencias, y a partir de ésta se deben representar los datos por medio de un diagrama de líneas y uno de barras, pero en la conclusión solo atienden a que con ayuda de éstas gráficas es posible entender mejor la información, pero no hay una explicación de cómo se hallan éstas.

Para la práctica de éste tema se dan los datos, con el fin de que los estudiantes construyan tanto la tabla de frecuencias como los diagramas de barras y líneas, también se trabaja con diagramas que deben ser analizados por los estudiantes, trabajo que resulta interesante. Y para el caso del trabajo en grupo, se trabaja como en el caso anterior, haciendo la encuesta, recogiendo datos, y construyendo la tabla de frecuencias y las correspondientes gráficas.

Luego vienen los diagramas circulares, en primer lugar se encuentra para lo que sirve un diagrama de éste tipo, se continua con el ejemplo y luego la conclusión en la cual se explica el proceso para hallar el valor en la circunferencia de cierta frecuencia, en el segundo ejemplo hay un pequeño error de escritura, que como se dijo anteriormente tal vez no es de gran incidencia en el proceso, pero puede llegar a generar desconfianza por parte del estudiante en su libro de texto. Al igual que en la anterior práctica se trabaja con tablas a partir de las cuales se pide hallar la correspondiente gráfica, se muestra el diagrama circular y se pide elaborar la tabla de frecuencias respectiva, caso para el cual el estudiante no tiene que aplicar el mismo proceso, si no que él esta en la posibilidad de hallar cuál sería la operación indicada para tal fin, y para terminar se tiene una situación problema, en la cual es necesario el análisis de un diagrama circular. Para el caso del trabajo en grupo se sigue trabajando en el contexto del curso, pero en este caso con el diagrama circular. Con el pictograma el trabajo es similar que se ha realizado con las dos anteriores representaciones, así que sería redundante describirlo, lo que resulta interesante es la variedad de diagramas y

representaciones a las que se refieren, las cuales son agradables de entender por parte del estudiante. Se continúa con las medidas de tendencia central, para las cuales se lleva a cabo un trabajo representativo y explicativo.

Luego se tiene lo que es la frecuencia relativa, trabajo para el cual se toma el mismo ejemplo de la temperatura utilizado en el desarrollo de la frecuencia absoluta, y se halla la frecuencia relativa de dichos datos, en la conclusión se explica como se halla, pero no dan la utilidad que éste puede tener a la hora de realizar un análisis de los datos, en el segundo ejemplo es posible encontrar las tres formas en las que se puede escribir la frecuencia relativa, fracción, decimal, y porcentaje.

En la práctica del trabajo con la frecuencia absoluta se encuentran actividades de completar tablas con las tres representaciones de la frecuencia relativa dada la frecuencia absoluta y el total de los datos, o también dado uno de los datos, ya sea la frecuencia absoluta, o la frecuencia relativa en fracción, decimal o porcentaje en dicha tabla, además se dan grupos de datos con el fin de organizarlos y hallar las respectivas frecuencias. En el trabajo de grupo con los datos obtenidos en las anteriores prácticas se trabaja con la frecuencia relativa.

La siguiente sección recibe el nombre de “Frecuencia acumulada I”. El estudiante se preguntará porque I, pues si la siguiente sección se titula “Frecuencia acumulada II”, entonces faltaría aclarar que la I se refiere a la frecuencia absoluta acumulada, y la II a la frecuencia absoluta relativa. Nunca se hace dicha aclaración, durante el desarrollo de los temas se habla simplemente de frecuencia acumulada, lo que no es de gran ayuda para que el estudiante luego trabaje con otros textos, y otras actividades, diferentes al texto guía.

Para el trabajo de la práctica en ambos casos se tienen tablas de frecuencias o diagramas de barras, a partir de las cuales se pide hallar las frecuencias acumuladas. En el trabajo en grupo se realiza lo mismo que se ha venido trabajando en las últimas secciones, pero con frecuencias relativas. En las actividades de nivelación hay preguntas sobre todos los temas,

pero son similares a las trabajadas en cada una de las prácticas, en el caso de las situaciones problemas, se encuentran problemas para los cuales se aplica todo lo visto durante el capítulo, pero en general el estudiante no está siendo incentivado a pensar, ya que se tienen preguntas que lo llevan a un procedimiento mecánico para su solución.

El nuevo examen de estado presenta cuestiones un poco más interesantes, ya que no todo está dado, el estudiante está en la oportunidad de ver nuevas cosas, y pensar un poco más, ya que en éstas cuestiones no se aplica simplemente un algoritmo. Las preguntas están divididas por competencias, como son la interpretativa, la propositiva e interpretativa, y sin duda cada uno de los grupos de actividades cumple con el fin de éstas. Al final, se presenta un resumen del capítulo, con algunas de las definiciones, o más bien con las instrucciones para hallar ciertas medidas estadísticas o frecuencias, se presenta el problema de no diferenciar la frecuencia absoluta acumulada y la frecuencia relativa acumulada.

#### **4.3.2. Grado 7º**

Los logros e indicadores de logro correspondientes para este grado en estadística son:

##### **En conocimiento matemático:**

- Identifica las características de una tabla de distribución de frecuencias:
- Determina la frecuencia relativa de una clase en una tabla de distribución de frecuencias

##### **En situaciones - problemas:**

- Analiza los conceptos de media, moda y mediana como medidas de tendencia central y las determina en un conjunto de datos.
- Determina la media, la moda, la mediana en un conjunto de datos.

##### **En comunicación matemática:**

- Representa de manera gráfica un conjunto de datos.
- Realiza un diagrama de barras o de líneas para representar un conjunto de datos.
- Distingue entre frecuencia absoluta y frecuencia acumulada
- Interpreta los datos de una tabla estadística y elabora conclusiones

Al leer los logros, podría parecer que el trabajo es el mismo que en texto del anterior nivel, esto no es así del todo, ya que en este caso las variables se trabajarán con intervalos, porque estas tendrán una clasificación, pero sin duda existen aspectos que si se repiten, como suele suceder con el trabajo en estadística en la secundaria.

Para el caso de la historia, se pasa al siglo XVII Allí se hace una breve explicación sobre la publicación del libro sobre los registros de mortalidad escrito por John Graunt, en el cual se establecieron grandes resultados como fueron halar el porcentaje de muertes por accidentes, suicidios, y algunas enfermedades, y también demostró que había más nacimiento de hombres que de mujeres.

En los requisitos se encuentran problemas, en los cuales se da la media o el promedio y se pide hallar los respectivos números que cumplen la propiedad, así pues que los estudiantes tendrán que aplicar una ecuación, y además manejar el concepto de la media.

Los temas que se trabajan en esta unidad son variables estadísticas, la distribución de frecuencias, la media y la clase modal y la mediana. El primer tema que se trabaja es el de variables estadísticas, se presenta un ejemplo en el cual es posible encontrar dos encuestas, una en la cual se hacen preguntas que tienen que ver con cualidades, y gustos, y en la otra preguntas que tienen que ver con medidas numéricas, luego se hacen diversas preguntas alrededor de éstas. Para terminar en la conclusión se refiere a lo que es una variable, pero no en estadística sino en matemáticas ya que la definen de la siguiente manera:

*Para representar los atributos o valores de cada uno de los elementos de un conjunto de datos empleamos una letra llamada variable. Por ejemplo, la variable X puede representar cada uno de los elementos de un conjunto {verde, azul, negro, café} y la variable Z, a las del conjunto {0, 1, 2, 3,4, 5, 6}.*

Trabajo que podría confundir al estudiante, ya que no siempre las variables van a ser las que se encuentran en los conjuntos, la explicación debería ser un poco más general. Luego de la definición explican cuando la variable es cualitativa y cuando es cuantitativa.



En el trabajo de la práctica, se pide clasificar un grupo de variables en cuantitativas y cualitativas, además se da una breve explicación acerca de lo que es una variable continua y una discreta, y de esta forma piden clasificar variables en continuas o discretas. Según la explicación que dan, las variables continuas pueden ser decimales, y la variable discreta solo pueden ser números enteros, lo que es falso. En el trabajo en grupo se pide hallar variables cuantitativas y cualitativas diferentes a las que se han nombrado durante el trabajo que se ha llevado hasta ahora, y se presenta un problema de análisis, pero no tiene nada que ver con las variables, es de razonamiento lógico. La segunda sección es acerca de distribución de frecuencias, se presentan los datos acerca de los puntajes que se obtuvieron en un concurso de ciencias, y se hacen diversas preguntas acerca de las frecuencias de pares de puntajes, se da la *tabla 4.9.*:

TABLA 4.9. SECCIÓN 2 PRENTICE HALL 7

Pedro	7	Cecilia	7	Elizabeth	8	Martha	7	Roberto	5
Juan	8	Luis	6	Andrés	10	Julio	9	Rodrigo	8
Marina	9	Bernardo	5	David	6	Jairo	7	Patricia	7
Santiago	5	Paula	5	Alex	7	Tatiana	7	Mónica	8

Se hacen preguntas como ¿Cuántos estudiantes obtuvieron puntaje de 5 o 6 puntos? ¿Cuántos de 7 u 8? ¿Cuántos de 9 o 12?, y luego se organizan dichos datos en la *tabla 4.10.*

TABLA 4.10. SECCIÓN 2 - PRENTICE HALL 7

Puntos	Número de alumnos y alumnas
5 o 6	7
7 u 8	12
9 o 10	3
Total	22

Y luego se dice que se tienen tres clases de puntajes, y que la tabla recibe el nombre de tabla de distribución de frecuencias, tal que en la primera columna muestra las clases en las que se agruparon los datos, y la segunda la frecuencia de clase.

Sería de gran ayuda hacer una comparación con lo visto en grado sexto, acerca de la frecuencia absoluta, lo que se está haciendo en este caso es lo mismo, la única diferencia es que se están tomando por grupos las variables. En este caso tomaron el ancho del intervalo de clase de dos, pero podrían mostrar una generalidad para éste.

Luego se halla el porcentaje de cada una de las frecuencias, respecto al número total de estudiantes, y se dice que éste valor recibe el nombre de frecuencia relativa, y la respectiva tabla de tabla de frecuencias relativas (*tabla 4.11*).

TABLA 4.11.

SECCIÓN 2 - PRENTICE HALL 7

Puntos	Número de alumnos y alumnas	Frecuencia Relativa
5 o 6	7	0.318
7 u 8	12	0.545
9 o 10	3	0.136
Total	22	1.000

Pero como se puede ver, en la tabla no se dan los valores del porcentaje, sino el decimal resultante al dividir cada una de las frecuencias entre el número total de los datos, así pues puede presentarse confusión entre el proceso que se realizó para hallar los porcentajes y la tabla. En la conclusión explican como hallar la frecuencia relativa de cada clase.

En la práctica se trabaja con tablas de frecuencias absolutas a partir de las cuales se pide hallar frecuencias relativas, frecuencias acumuladas, y diagramas de barras, temas que fueron vistos en el grado anterior, también hay una tabla con cierta información a partir de la cual se plantean ciertas preguntas de análisis y lectura. Para el trabajo en grupo se pide hacer una encuesta en el curso, a partir de la cual se pide hallar una tabla de frecuencias absolutas, relativas, acumuladas y diagrama de barras.

Las actividades de nivelación, las situaciones problemas y el nuevo examen de estado presentan en general el mismo esquema del grado anterior, está demás referirse a estos.

### **4.3.3. Grado 9º**

El título de la unidad no es simplemente estadística, es Estadística y Probabilidad y los logros e indicadores de logro correspondientes para este grado en el área de estadística son:

#### **En conocimiento matemático:**

- Diferencia un dato cuantitativo de uno cualitativo y un atributo de una variable.
- Identifica los datos que se requieren para un estudio estadístico.
- Clasifica los datos estadísticos de acuerdo con la observación de las variables de una muestra e identifica la población y la muestra.
- Calcula las principales medidas estadísticas; halla la mediana y la media de un conjunto de datos.

#### **En comunicación matemática:**

- Distingue las situaciones en las que se emplean las probabilidades
- Organiza datos estadísticos

#### **En situaciones - problemas:**

- Resuelve problemas de probabilidad
- Calcula la probabilidad de un evento aplicando la regla de la suma

Analizando los logros e indicadores, se puede observar que se va a seguir trabajando con los estadígrafos de descripción, y las variables cuantitativas y cualitativas, de lo cual no hay tanto que decir, pero ellos consideran importante seguir trabajando con lo mismo, en vez de avanzar un poco en las temáticas y trabajan temas que debieron ser tratados en grados anteriores como son la definición de población y muestra. Además se empieza a trabajar con las probabilidades pero solo las toman en situaciones problemas y comunicación, no se tiene en cuenta el conocimiento matemático que se debe tener acerca de éstas.

En la parte histórica se tiene el siglo XVI, época durante la cual las estadísticas empezaron a progresar y se convirtieron en uno de los instrumentos de medición y análisis empleados como soporte de la estructura administrativa de las monarquías, es así que se escriben las

primeras obras de estadística con un carácter más descriptivo que analítico. A mediados del siglo XVII se da un salto a la aplicación a otras disciplinas que eran hasta entonces puramente descriptivas, como son la economía, la demografía, y las ciencias sociales y se progresa en el cálculo de las probabilidades, que había surgido tiempo atrás como una necesidad para interpretar juegos de suerte y azar.

En los requisitos, se presentan ciertas afirmaciones, acerca de falacias estadísticas, que pueden ser una trampa para el lector, a la hora de publicarlas en un periódico u otro medio masivo de comunicación, y se recomienda como ejercicio leer la prensa de la semana en la que se encuentre el estudiante y buscar noticias que le induzcan a pensar que está en frente de una falacia estadística. Los temas a trabajar en este nivel son, la población y muestra, las variables estadísticas, las observaciones y datos estadísticos, la organización y presentación de datos cualitativos, la construcción de una distribución de frecuencias, el histograma y polígono de frecuencias, la media aritmética, la media aritmética de una distribución de frecuencias, la media ponderada, la mediana, los percentiles, la moda, la probabilidad, experimentos aleatorios y espacios muestrales, el calculo de la probabilidad de un evento, la probabilidad de la unión de eventos.

Como se puede notar muchos de los temas que en este grado se trabajan ya han sido trabajados en anteriores niveles, así pues que este espacio podría ser utilizado para estudiar y avanzar en el área de la probabilidad y la estadística.

En la primera sección se trabaja dando la diferencia entre estadísticas y estadística, para esto muestran ejemplos de cada una de ellas y al final en cada una de las conclusiones explicitan lo que son. *La agrupación de cifras relacionadas con un aspecto en particular se denominan estadísticas. El conjunto de métodos que se emplean para recoger, procesar, analizar e interpretar datos con el fin de obtener conclusiones útiles a partir de ellos se denomina estadística.*

En el trabajo de la práctica se realizan afirmaciones acerca de estadísticas y estadística sobre las cuales se debe dar una opinión sobre su veracidad, trabajo que considero subjetivo, podría ser útil para crear una discusión en el salón de clase. Para el trabajo en grupo se pide buscar en publicación estadística de los últimos 10 años acerca del valor del

salario mínimo, valor y detalle de las exportaciones de café a países europeos, producto interno bruto, índice de precios al consumidor, y tasa de desempleo. En mi opinión el trabajo resulta interesante ya que además de la realización de una investigación en diversas fuentes, el estudiante tendrá la oportunidad de conocer sobre temas que pueden ser de su interés y además les pueden ser útiles en algún momento de su vida.

La parte dos, se titula población y muestra, se da un ejemplo en el cual se determinan los conjuntos de estudio de los ejemplos de la sección anterior, estos recibirán el nombre de población o universo, y lo definen como el conjunto de personas y objetos cuyas propiedades se han de analizar. Y para la muestra dan un ejemplo de un estudio, y cual sería la muestra de la población en ese caso, así pues la muestra la definen como un subconjunto de la población o universo en estudio. Se continúa con las variables estadísticas, y se trabaja al igual que en grado 7° con las cualitativas y cuantitativas, presentando un ejemplo en el cual se tiene un objeto que posee diversas características, como son el color, el tiempo de uso, el estado, el precio, el tamaño y el número de velocidades, logrando así hallar de ambos tipos de variables en éste. El elemento adicional que aparece en este grado es la definición de lo que es un atributo, aquellos elementos que no pueden medirse, es decir las variables cualitativas son atributos. Cuando se refieren a medida deberían hacer énfasis a que están hablando de medida numérica.

En el trabajo de la práctica se dan ciertos ejemplos de variables y se pide clasificarlas en cualitativas o cuantitativas, se pregunta acerca de la variabilidad acerca de conjuntos de datos, y se dan situaciones en las cuales se pide señalar las variables que serán estudiadas. Para el trabajo en grupo se pide identificar las variables cualitativas y cuantitativas que puede tener un estudiante típico del colegio. La sección 4 se titula observaciones o datos estadísticos, la inician diciendo que éstos son la materia prima sobre la cual trabaja la estadística, en el ejemplo se hace un cuestionario con el fin de recolectar datos acerca de tres estudiantes con preguntas del estilo: ¿Cuál es su sexo? ¿Cuál es su edad (en años cumplidos)? ¿Cuál es su peso (en kilogramos)?, etc. y se obtiene la tabla 4.12. de resultados:

TABLA 4.12. SECCIÓN 4- PRENTICE HALL 9

Persona	Respuestas a las preguntas de la encuesta					
	1	2	3	4	5	6
1	M	16	55,4	Negro	No	D
2	F	15	50,7	Café	No	D
3	M	15	53,8	Castaño	Si	Z

Definiendo luego lo que es un dato estadístico, dato cualitativo o categórico, y dato cuantitativo o numérico. Este tema debió haber sido visto en el grado sexto, es decir en el inicio del estudio de la estadística, además podrían hacer una mejor clasificación de las variables, por ejemplo dependiendo de su escala.

En la práctica se encuentran situaciones en las cuales hay que clasificar los datos como cualitativos o cuantitativos, como vemos vuelve y se reitera un tema que no es tan complicado, pero sin embargo ha sido tomado durante dos niveles mas. En la parte cinco se trabaja con la organización y presentación de datos cualitativos, en primer lugar se encuentra un tabla acerca del sabor de gaseosa que prefieren 40 estudiantes de grado séptimo, luego afirman que solo hay dos formas de representar los datos cualitativos, un histograma y un diagrama de sectores, lo que resulta ser un desacierto ya que existen muchas mas formas de representación, como son el diagrama de tallo y hoja, el pictograma y otros más, además el histograma es utilizado para representar los datos cuantitativos continuos, por la continuidad de las variables en el eje x, y las barras pegadas.

Para la práctica se dan los datos de la evaluación de 36 estudiantes de un curso sobre el manejo de las leyes que rigen las operaciones entre conjuntos y se pide elaborar una tabla, y los diagramas correspondientes, también se tiene el resultado de una estadística de la cual se pide hallar las gráficas y analizarlas con el fin de dar una conclusión acerca de los resultados de la encuesta. En el trabajo en grupo la idea es hacer tres preguntas a los compañeros de curso y hacer un completo informe estadístico sobre los datos que se recojan.

En la parte 6 se realiza la construcción de una distribución de frecuencias, este trabajo ya se realizó en el grado anterior, pero en este caso ya es más general ya que se puede encontrar una manera general para hallar los intervalos de clase, ya sea para variables discretas o continuas, además dan los respectivos procedimientos para hallar la marca de clase, las frecuencias absolutas, y las frecuencias relativas, pero no muestran cual puede ser el uso de éstas. La distribución de frecuencias no es un concepto, es un trabajo que se lleva a cabo para lograr un mejor análisis de las variables con las cuales que se está trabajando.

En la práctica se dan datos, y se pide construir las correspondientes distribuciones de frecuencias, y en algunos casos representarlos por medio de un histograma o un polígono de frecuencias, pero ¿dónde se explicó el procedimiento para graficar un polígono de frecuencias?

Para la sección siete se trabaja con histogramas y polígonos de frecuencias, se tiene una tabla con variables divididas en intervalos de clase, la frecuencia correspondiente a cada una de ellas, el porcentaje correspondiente, y los puntos medios de los intervalos, y su representación por medio del un histograma y un polígono de frecuencias. Luego afirman que una distribución de frecuencias se representa gráficamente por medio de un histograma y un polígono de frecuencias, olvidando que hay muchas más formas de representación. Además la tabla podría tener más datos acerca de las distribuciones acumuladas.

En las conclusiones se explica cual es el proceso para graficar un histograma y un polígono de frecuencias, en este caso, no toman los puntos medios de las barras del histograma para explicar la representación por medio del polígono de frecuencias, procedimiento que podría resultar más asequible para el estudiante. En la práctica se tienen tablas con datos y partir de éstas se pide representar los datos por medio de un histograma y un polígono de frecuencias, podría haber un ejercicio de análisis de gráficas. Desde la sección ocho hasta la trece se trabaja con estadígrafos de descripción, y del la catorce hasta la dieciséis con probabilidad, experimentos aleatorios, espacios muestrales, calculo de probabilidad de un

evento y probabilidad e la unión de eventos. Entran de nuevo las medidas de tendencia central, pero en todos los textos se emplean diferentes ejemplos y así mismo distintas formas para hallar dichos valores, y para el caso de los temas de probabilidad en mi opinión es muy poco espacio para la explicación de éstos temas que van a jugar un papel importante durante el desarrollo de toda la estadística inferencial.

#### **4.3.4. Grado 10°**

El título de la unidad ya no es estadística, ni Estadística y Probabilidad es Probabilidad los logros e indicadores de logro correspondientes para este grado en el área de estadística son:

##### **En conocimiento matemático:**

- Enumera y describe las principales técnicas de conteo.
- Realiza diversas combinaciones de situaciones dadas.
- Define probabilidad y construye un concepto de probabilidad con el enunciado de varios eventos.
- Explica el significado de los conceptos básicos de las probabilidades.
- Refiere diversos sucesos en los que reconoce los elementos básicos.
- Describe los enfoques de la probabilidad clásica y de frecuencia relativa.
- Explica las diferencias de los enfoques de la probabilidad clásica y frecuencial.

##### **En situaciones - problemas:**

- Calcula probabilidades para situaciones sencillas y de diversos sucesos.
- Emplea las reglas de la suma y complementación para el cálculo de probabilidades.
- Diferencia eventos excluyentes de los que no son mutuamente excluyentes.
- Calcula probabilidades en condiciones de dependencia e independencia.
- Distingue eventos dependientes e independientes con base en diversos sucesos.

##### **En comunicación matemática:**

- Explica que es probabilidad condicional y probabilidad independiente
- Clasifica en condicional e independiente un evento



En este libro se trata más a fondo todo lo que tiene que ver es probabilidad básica, así pues que se ve la innecesidad de incluir algunos de los temas en grado 9º, ya que aquí se encuentran explicados más a fondo cada uno de los temas. En la conexión histórica se tienen trabajos desde el siglo XVII hasta el siglo XVIII, los cuales tratan principalmente acerca de juegos de azar y un libro que escribió Cardan acerca de este tema, en el cual se encuentra implícita la ley de los grandes números así como también el cálculo de probabilidades para obtener algunos resultados en juegos de cartas y especialmente en el póquer medieval, el cual fue escrito alrededor del año 1520 y publicado en 1663. El inicio formal de la probabilidad se da con Pascal (1623–1662) y Fermat (1601– 1665) cuando tratan de dar soluciones a problemas relacionados con juegos de azar, más adelante otros matemáticos hicieron grandes aportes a ésta área como fueron Moivre (1667-1754), Laplace (1749-1827) y Gauss (1777-1855). En el trabajo con requisitos, lo que se hace es trabajar con permutaciones y combinaciones, lo cual tal vez ha sido visto por los estudiantes en otros cursos. No se adentrará en ninguno de los temas, ya que no son relevantes para el desarrollo de la unidad didáctica.

Dentro del análisis de éstos dos textos cabe destacar el tratamiento que se da al inicio de cada capítulo acerca de diferentes aspectos como son el ámbito histórico, las diversas aplicaciones y una prueba de entrada, aspectos que sin duda logran ser interesantes para los estudiantes. Sin embargo ¿Por qué no usar esos campos de aplicación dados en esta introducción, como contextos en problemas planteados para que los estudiantes desarrollen en el transcurso del capítulo?

Como no hay contextos adecuados el trabajo se torna repetitivo, ya que los temas en que están basados los problemas, se refieren en general a supuestas características observadas dentro de un salón de clase, en ocasiones utilizadas sin sentido alguno, o con cierta discordancia al ser comparadas con la realidad; este aspecto genera una notable falta de interés en el educando, ya que él siempre se encuentra en una continúa búsqueda para responderse a sí mismo ¿Para qué me sirve?

Si una persona no le encuentra utilidad a un trabajo que desarrolla simple y llanamente, pierde el interés, herramienta fundamental para lograr los objetivos propuestos dentro de la enseñanza.

En referencia al desarrollo de los capítulos, en estos no se nota un trabajo de construcción, al contrario se introducen definiciones sin ninguna clase motivación, o se dan datos sin una explicación, porque fueron tomados, o sobre que población se hizo el estudio, además el estudiante es privado de tomar sus muestras para que él a partir de éstas llegue a construir determinados conceptos.

Otro aspecto que es importante tener en cuenta es que no hay espacios abiertos para que el estudiante conjeture, todo le es dado, a él no se le es la posibilidad de hallar conjeturas acerca de un conjunto de datos observados, criterio que resulta fundamental en cierto análisis estadístico.

#### 4.4. Resumen de temas considerados en las colecciones analizadas

Los dos colecciones de libros analizados en sus diferentes niveles son mostrados en la siguiente tabla:

TABLA 4.13  
RESUMEN CONTENIDO TEXTOS ANALIZADOS

TEXTO TEMA	ALFA						PRENTICE HALL					
	6	7	8	9	10	11	6	7	8	9	10	11
Análisis de datos	x						x			x		
Clases de Variables		x						x		x		
Tablas de Frecuencias		x					x			x		
Distribución de Frecuencias										x		
Frecuencias absolutas	x						x	x				
Frecuencias relativas	x						x	x				
Frecuencia acumulada absoluta	x						x					
Frecuencia acumulada relativa	x						x					
Moda, mediana, media aritmética	x	x	x				x	x		x		
Media ponderada										x		
Percentiles										x		
Clase modal			x									
Clase mediana			x									
Clase para la media			x							x		
Diagramas de barras	x						x					
Diagrama de Líneas							x					
Histograma		x	x							x		
Polígono de frecuencias			x							x		
Diagrama circular	x						x			x		
Pictograma	x						x					
Subconjuntos		x										
Principio fundamental conteo					x						x	
Permutaciones y combinaciones		x			x						x	
Sucesos aleatorios		x	x							x	x	
Espacios muestrales					x					x	x	
Concepto de probabilidad		x	x		x					x	x	
Postulados de probabilidades			x		x					x	x	
Variables aleatorias						x						
Valor Esperado						x						
Varianza						x						
Desviación Estándar						x						
Distribución binomial						x						
Distribución de Poisson						x						
Distribución normal						x						

## 4.5. Contextos

En MEN (1985) se afirma que la toma de decisiones es de gran importancia para la diversificación de procesos empresariales, el aumento del volumen en la producción de los sectores de la actividad económica, la especialización en distintas áreas del conocimiento, la diversidad de servicios generados en los sectores sociales, y los avances en el campo tecnológico. Además, hace unos cien años han empezado a utilizarse los métodos de la estadística como una poderosa técnica en la investigación científica experimental y más recientemente en la llamada investigación operativa, aplicada en la industria, problemas de gobierno, negociaciones de alto nivel, estrategia militar y otras ramas. Confiriendo así un lugar importante a la estadística como instrumento de análisis y predicción en el desarrollo social y económico de las sociedades. Es así que juega un papel de suma importancia en el desarrollo económico, social y político de cualquier país.

Además es necesario tener en cuenta que la estadística no es un conocimiento demasiado especializado, sino que sin duda afecta la actividad cotidiana de todo hombre que esté en contacto con los medios de comunicación que lo rodean. Estos hechos hacen que deban considerarse hoy las ideas estadísticas básicas como una parte indispensable en la mente del hombre educado y el método estadístico en conjunto como un instrumento de trabajo esencial para el economista, el ingeniero, el demógrafo, o para el investigador experimental (biólogo, químico, médico, psicólogo...) y para el analista operacional en el campo militar, económico o industrial. A continuación se nombraran algunos ejemplos dados por García (1995) de la vida diaria en los que se vislumbra la necesidad de la estadística:

- A pesar de la idea de porcentaje, es corriente encontrar noticias de prensa como ésta: en la cual que al aumentar la producción diaria de una fábrica en 2.5%, el aumento semanal (cinco días) resulta ser 12.5%
- La idea de promedio es sencilla, sin embargo resulta curioso observar que muy pocas personas dan la respuesta correcta a la siguiente pregunta: “Si usted conduce un automóvil una distancia de 20 Km. a una velocidad de 80 Km/h. y después otros 20 Km. a 90 Km/h., ¿Cuál es la velocidad media?

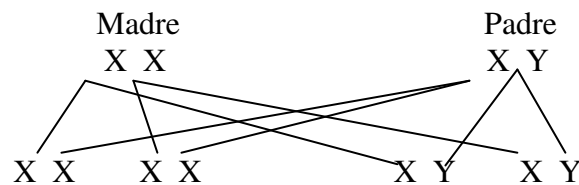
- Se dice frecuentemente en la prensa que la proporción de muertos por cáncer aumenta constantemente, pero para que tal afirmación tenga un sentido claro y no pueda interpretarse como un fracaso de la medicina en tal tratamiento, deben tenerse en cuenta, factores como: el diagnóstico es actualmente más preciso, de modo que muchos casos que hace unos años se habrían atribuido a causa desconocida se incluyen hoy en el concepto del cáncer; la recopilación de datos estadísticos es hoy más completa que hace años, e incluso se practican en algunos países diagnósticos post mortem de fallecidos por causa desconocida; el progreso extraordinario en la curación de otras enfermedades y la higiene han dado como resultado una disminución en la mortalidad infantil; el cáncer es una enfermedad principalmente de viejos y no es raro que haya aumentado la correspondiente tasa de mortalidad.
- La seguridad relativa a los viajes aéreos es un tema de gran actualidad y sobre el que se hace mucha propaganda periodística. Los que desean demostrar que el riesgo es muy pequeño, hacen estadísticas en que figura el número de accidentes por pasajero y kilómetro recorrido. En cambio los que quieren exagerar el riesgo presentan las estadísticas de accidentes por pasajero y hora de vuelo. Tanto uno como otro punto de vista son excesivamente simplistas, ya que es muy diferente el riesgo de un vuelo según sea con o sin escalas, por encima del mar o no, etcétera.
- La consideración abusiva del promedio sin tener en cuenta la dispersión, que fue en otros tiempos grave defecto de la estadística, puede hacerse patente con el siguiente problema interesante para los automovilistas: Es bien conocido que el consumo de gasolina depende de la velocidad del automóvil, disminuye al aumentar la velocidad hasta un mínimo y luego vuelve a aumentar con la velocidad. Si la velocidad de consumo mínimo es 50 Km/h, el que un conductor haya tenido esta velocidad media durante un largo viaje no permitirá afirmar de modo alguno que habrá hecho el mínimo consumo.

Haciendo referencia ahora al tema del azar, el cual es un fenómeno que también rodea el individuo, ya que puede éste ser observado a través de muchos aspectos del entorno, así como por medio de la realización de juegos y actividades, que pueden llegar a ser asequibles por el maestro y el estudiante dentro del aula de clase.

A continuación se nombraran algunos fenómenos aleatorios que pueden ser utilizados para la ejecución de situaciones didácticas, que contengan estadística o probabilidad. Se van a nombrar los cuatro grandes grupos que se describen Tanur y cols (1971) para clasificar los campos de aplicación estadística: el hombre y su mundo biológico, físico, social y político.

#### 4.5.1. Biología

En este campo es posible hablar del hecho de que muchas de las características heredadas como son el peso al nacer, el sexo o el color de pelo no se pueden prever, éstas dependen del azar. Un ejemplo de que la transmisión de caracteres genéticos obedece a las leyes de la probabilidad, es el caso del sexo de un individuo, ya que el sexo femenino contiene el para de cromosomas XX, y el sexo masculino XY, es así que la madre transmite un cromosoma X, y el padre puede enviar uno X o uno Y, de esta forma aproximadamente la mitad de los recién nacidos serán varones, veamos una diagrama de este fenómeno:



Además se podría hablar de otras aplicaciones en la medicina como son, la edad en que sufre cierta enfermedad, la duración de determinado síntoma, la posibilidad de un diagnóstico correcto cuando hay diversas enfermedades que presentan los mismos síntomas, la posibilidad de contagiarse de cierta epidemia, el efecto de una vacuna, el modo en que se determina el recuento de glóbulos rojos a partir de una muestra de sangre, o la posibilidad de heredar un defecto o una enfermedad. Para hacer estudios mundiales, ya sea sobre la extinción de algún ser vivo, para hacer estimaciones acerca de la extensión de una

cierta enfermedad o la esperanza de vida de un individuo se utilizan estudios probabilísticos de crecimiento de poblaciones.

Estos modelos en agricultura y zootecnia son utilizados para prever los efectos del uso de fertilizantes o pesticidas, o para evaluar el rendimiento de una cosecha. En el ámbito de la psicofisiología, se ve el efecto del azar en el cociente intelectual, en los tipos diferentes de caracteres o capacidad de los individuos o la intensidad de respuesta a un estímulo.

#### **4.5.2. Física**

Los fenómenos meteorológicos son un excelente ejemplo a la hora de hablar de aleatoriedad, como ejemplos de éstos se tienen las temperaturas máximas y mínimas, la intensidad y dirección del viento, las tormentas o granizos, y la duración, intensidad y extensión de las lluvias. Así como también las consecuencias de dichos fenómenos que tiene que ver con la magnitud de los daños que puede llegar a producir una granizada o como el agua de un lago o pantano puede llegar a aumentar su volumen a causa de la lluvia.

Otras variaciones que son de un carácter aleatorio, y pueden estar clasificadas dentro de ésta área, son las que están sujetas a estimación de necesidades, precio y localización de fuentes de energía de materias primas como lo son el carbón, el petróleo y otros minerales.

También es posible encontrar los problemas de fiabilidad y control de calidad de electrodomésticos o dispositivos que son usados por la comunidad, como son un televisor, una grabadora, o hasta el mismo carro.

Para terminar se encuentran las magnitudes, como son la longitud, el tiempo y el peso, las cuales resultan ser una fuente de variabilidad, ya que uno de los problemas que pueden ser planteados son los que se refieren a la estimación del error, a la hora de asignar el valor más preciso a determinada medida.

### **4.5.3. Social**

En las situaciones que vive a diario una persona esta presente la incertidumbre, desde la familia, hasta la escuela, el trabajo o el ocio, por ejemplo en la edad de dos personas al contraer matrimonio, o las hobbies de los integrantes de una familia, hasta el número de hijos de cada una de las familias.

En los juegos de ocio, se practican los juegos de azar, o cuando se adquiere una póliza de seguro no se sabe si se cobrará o se perderá el dinero pagado. En fin en cada situación que se vive a diario, están presentes los sucesos aleatorios.

### **4.5.4. Política**

A la hora de tomar cualquier decisión, el gobierno o cualquier organismo político está en la necesidad de investigar, lo cual es posible con la aplicación de censos y encuestas, las variables aleatorias que se investigan están basadas en dichas estadísticas basadas en la colectividad. Entre las más importantes se encuentran la producción y precio de bienes, el comercio, las estadísticas demográficas, la emigración e inmigración, etc., de las cuales es posible obtener información por medio de los diversos medios de comunicación.

Todas las anteriores aplicaciones pueden ser relacionadas con la vida diaria de los estudiantes, y de esta forma mostrarles la variedad de aplicaciones que tiene esta rama de la matemática.

## **4.6. Uso de tecnologías en el aula de clase**

Actualmente se piensa que las únicas tecnologías son las computadoras, las calculadoras graficadoras, los teléfonos celulares, etc.; sin imaginar que en su tiempo la escritura fue un gran invento que sirvió como amplificador del pensamiento humano, con esta invención, se podría decir, que se dio inicio a una comunicación más estable y permanente.



Es innegable que el hombre va evolucionando teniendo en cuenta sus necesidades, todo surge de acuerdo al tiempo y a las posibilidades, los objetos que hoy poseemos, desde una simple herramienta como lo es el lápiz, hasta un objeto de entretenimiento, como es la televisión, son el producto de una constante evolución desde nuestros antepasados quienes a partir del más mínimo elemento estuvieron en la capacidad de lograr un sinnúmero de inventos que fueron contribuyendo a la virtualidad.

Todas las herramientas son extensiones de las capacidades del hombre, no son hechas para hacer más fácil cierto trabajo, sino para sistematizarlo o perfeccionarlo, por esta razón son creadas por ciertos individuos, duplicadas por otros, y de gran uso para toda la humanidad en general; claro, sin llegar a imaginar que estos objetos podrían ser tan capaces de producir como lo es la inteligencia de un ser humano,

Para citar un ejemplo los computadores solo son un medio de materialización del pensamiento, en esta es posible introducir información, y sirve como instrumento para producir y socializar, aspecto que es de suma importancia para continuar el proceso de desarrollo, ya que todo lo que sea comunicado será en pro de la evolución, claro desde que esta información no sea utilizada para fines fuera maquiavélicos.

Esto quiere decir que si la evolución continua de la misma forma, lo que hoy para nosotros es tecnología de punta, serán simples herramientas para las generaciones futuras y así sucesivamente.

Entonces la idea con el presente trabajo es entrar a formar parte de esta gran comunidad evolutiva, con un programa que facilitará la incursión del estudiante dentro del análisis estadístico.

## CAPÍTULO 5

### UNIDAD DIDÁCTICA

#### **5.9. Ubicación en el currículo**

Según los textos estudiados, y los lineamientos curriculares, el tema de distribuciones de frecuencias está ubicado en el grado 7º, y las distribuciones de probabilidad en el grado 11º. Así pues que lo que se hará será reunir dos actividades, en donde cada una podrá ser aplicada a cada uno de los anteriormente descritos.

#### **5.10. Justificación**

Según Baena (1999) el hombre se comunica a través de la información, la cual no es innata en los seres, esta se genera en el momento en el cual el ser relaciona, discrimina, abstrae, y generaliza sus observaciones del mundo exterior. De esta forma es que surgen las primeras informaciones acerca de los objetos como lo son el nombre, el color el tamaño, etc., y así a medida que el hombre se va desarrollando va llevando un proceso de clasificación, y es en ese preciso momento donde se ve la importancia de implementar métodos para facilitar dicho proceso, en cierta etapa del desarrollo cognitivo, y el espacio es propiciado por el área de la estadística.

Es indudable que estamos viviendo una era en la cual el ritmo de producción de información y comunicación es muy veloz, diariamente el cerebro recibe tanta información que llega un punto en que resulta difícil asimilarla, es aquí donde las herramientas y conocimientos van a ser necesarios para conseguir determinadas informaciones o llegar a tomar decisiones de gran importancia. La mejor herramienta para sobrevivir en la sociedad es tener la mente bien organizada, y que este en la capacidad de generar, representar, almacenar, transmitir y acceder a la información. Es por esta razón que estas metas están incluidas dentro del currículo y es importante que el maestro esté en la capacidad de enseñarlas.

Existen diversas actividades con las cuales los estudiantes pueden plantar preguntas y resolver problemas que estén basados en la observación, búsqueda, selección, recogida, organización, registro, reducción e interpretación de datos, y es posible desde las matemáticas contribuir a que los estudiantes adquieran la capacidad de desenvolverse en esta sociedad de continuo avance en la cual estamos viviendo.

Como se ha estudiado anteriormente la mayoría del trabajo en el libro de texto consiste en a partir de datos o tablas dados de antemano, estudiar cuestiones como media moda, desviación típica, etc., con datos que los estudiantes no saben siquiera de donde provienen, es necesario que el estudiante empiece procesos de análisis de información dada, formulación de preguntas respecto a observaciones, definición de variables de las unidades de estudio, consulta de fuentes, interacción con sistemas de datos que estén dentro de su contexto, elaboración de sistemas de datos del ambiente que lo rodea, lo que implica recoger datos, diseñar y planificar como realizar dicho proceso. Todos estos y más aspectos serán los cimientos de una correcta formación estadística.

Así como el cálculo aritmético, algebraico o vectorial esta presente en la física, es necesario que la estadística tenga contacto con otras áreas, como lo son las ciencias de la naturaleza y las ciencias sociales, y todas aquellas áreas en las que sea necesario algún calculo estadístico para la resolución de algún problema.

Es así que la enseñanza de la estadística debería ser enfocada desde que los estudiantes formulen preguntas, elaboren instrumentos de recogida de datos, elaboren tablas y resúmenes, terminando con la interpretación de parámetros. La idea principal es establecer metodologías apropiadas para tal fin, lo cual podría ser posible haciendo que el estudiante esté conectado con sus intereses y el entorno en el cual se desarrollan, de esta forma Baena (1999) afirma que es necesario que los estudiantes recorran todas las fases desde la definición del problema hasta la redacción y presentación de conclusiones, como para este caso no se está estudiando toda la estadística, sino una pequeña parte de ésta, entonces se hará algo similar, pero con las distribuciones estadísticas y sus representaciones.

### **5.11. Prerrequisitos**

- Idea de los objetivos de un estudio estadístico.
- Concepto de población y muestra.
- Concepto de variable.
- Diferenciar entre la variable y su frecuencia.
- Reconocer los datos de un problema.
- Variables aleatorias
- Conceptos y propiedades de probabilidad
- Combinaciones y permutaciones

### **5.12. Capacidades que desarrollarán los estudiantes**

- Plantear preguntas que se resuelvan con un estudio estadístico.
- Determinar variables y categorías de algún problema
- Adquirir estrategias de recuento.
- Organizar y estructurar datos.
- Ordenar y clasificar datos.
- Elaborar e interpretar tablas de frecuencias.
- Diseñar encuestas.
- Obtener datos de forma individual y colectiva utilizando diferentes fuentes y recursos.
- Organizar y estructurar datos utilizando diferentes estrategias de recuento.
- Resolver problemas en los que sea necesario utilizar la estadística.
- Usar e interpretar críticamente las tablas y las gráficas estadísticas.
- Reconocer las distribuciones de tipo continuo y las de tipo discreto.
- Utilizar las distribuciones de probabilidad para solucionar problemas de incertidumbre.

## **5.13. Contenidos**

### **5.13.1. Conceptuales:**

- Clasificación de variables estadísticas.
- Distribuciones de Frecuencia.
- Representación de las distribuciones de frecuencia.
- Distribuciones de probabilidad continuas y discretas
- Representación de las distribuciones de probabilidad

### **5.13.2. Procedimentales:**

- Diferenciación de las variables cuantitativas de las cualitativas.
- Determinación de variables continuas y discretas.
- Reconocimiento de escalas en nominales, ordinales, de intervalo y de razón.
- Interpretación de problemas de tipo estadístico.
- Utilización de tablas y gráficas dependiendo el contexto.
- Análisis de problemas por medio de la distribución de frecuencias y sus diversas representaciones.
- Interpretación de gráficas y tablas.
- Resolución de un problema probabilístico por medio de distribuciones de probabilidad.
- Interpretación de las diversas representaciones de las distribuciones de probabilidad.

### **5.5.3. Actitudinales:**

- Apreciar la utilidad del lenguaje gráfico para la representación de situaciones diversas
- Subrayar la variedad de situaciones en las cuales se puede aplicar la estadística y que son utilizados diariamente.
- Desarrollar una actitud crítica ante la presencia de información estadística.

- Mostrar interés y curiosidad para plantear encuestas y clasificar información por medio de tablas o gráficas de frecuencias.
- Destacar la necesidad de diversificar las gráficas dependiendo el contexto y la información obtenida.
- Reconocer la información de un problema para aplicar una distribución de probabilidad ya sea continua o discreta.

## **5.6. Metodología**

Se propondrán una serie de actividades de dos tipos, la primera serie de actividades considera contextos reales, para la adquisición de nuevos conocimientos; dentro de éstas se podrá encontrar algún trabajo de campo, en el cuál el estudiante tenga que llevar a cabo encuestas y estar en contacto con la comunidad en la que se desarrolla. Estas actividades se pueden trabajar de forma individual o en grupo. Además, algunas de ellas estarán ya resueltas para que los educandos adquieran estrategias propias de la actividad matemática y estadística, y tengan modelos que mejoren su modo de presentación, planteamiento y resolución de los problemas.

En el caso de la segunda serie de actividades se propone la utilización del software Minitools, el cual está dividido en tres fases fundamentalmente, en la primera de ellas el estudiante va a iniciar su trabajo identificando frecuencias de datos por medio de análisis de gráficas, las cuales serán introducidas utilizando una situación problema diseñada por el maestro, en la segunda parte ya tendrá la posibilidad de crear grupos de datos y así llegar a construir algunos de los conceptos y gráficas fundamentales para el estudio de la distribución de frecuencias y para terminar se tendrá la posibilidad de establecer conjeturas más exactas acerca de los datos por las múltiples herramientas que allí se presentan.

Para terminar la idea es llevar a cabo una puesta en común del trabajo realizado, en ambas fases analizando y debatiendo sobre los conceptos y materiales utilizados, las dificultades encontradas y las soluciones adoptadas para soslayarlas, esta actividad en mi opinión

resulta siendo una de las más importantes para lograr una retroalimentación del trabajo hecho hasta el momento.

En algunos casos, solo en los que estrictamente necesario se llevará a cabo una exposición detallada de los conceptos que aparecen en la realización de las actividades, por parte del profesor, intentando no perder el contacto con el estudiante.

## **5.7. Actividades del primer tipo (Trabajo en el aula)**

En primer lugar resulta necesario que el estudiante se familiarice con las distintas clases de variables, y escalas, así pues que será necesario un trabajo en clase, y otro de campo para observar si se ha entendido el tema.

### **5.7.1. Tarea 1. Variables cuantitativas vs. variables cualitativas**

En esta primera actividad se e presentaran diversas informaciones al estudiante por medio de tablas y gráficas, además de que se irá familiarizando con dichas representaciones, podrá ver algunas características y tal vez llegar a alguna clasificación de variables:

Por ejemplo se tienen, las siguientes dos gráficas, tomadas de la revista semana, la primera de ellas tiene que ver con los ingresos operacionales de ciertas empresas que manejan el Internet en el país, los cuales están representados por medio de un diagrama de barras, la idea es hacer una observación, y si el estudiante necesita ayuda, responder las preguntas necesarias solo respecto a la gráfica.

La segunda se refiere a las reacciones respecto al tema de la reelección del presidente Uribe electo en 2005, y se hace una representación por medio de un diagrama circular.

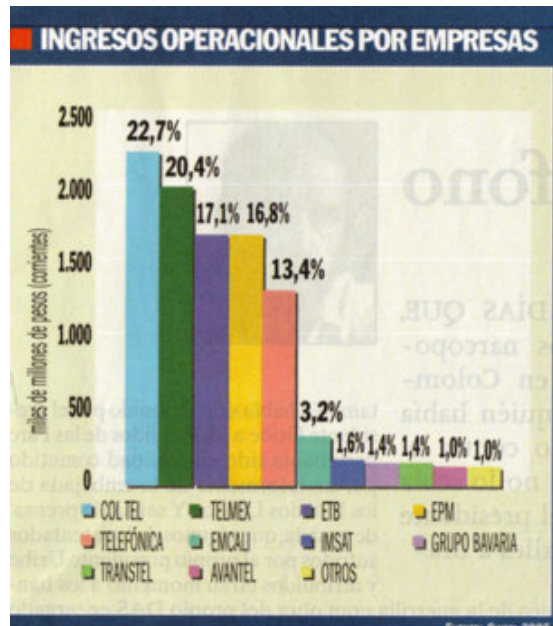
A continuación se le pide al estudiante que responda las siguientes preguntas respecto a ellas, y de esta forma podrá llegar a distinguir los diversos tipos de variables:

GRÁFICA 5.1.  
REACCIONES DE FALLO ACERCA DE LA REELECCIÓN DEL PRESIDENTE URIBE ELECTO AÑO 2005



FUENTE: REVISTA SEMANA DEL 24 AL 31 DE OCTUBRE DE 2005 PAGINA 41

GRAFICA 5.2.  
INGRESOS OPERACIONALES DE ALGUNAS EMPRESAS DE INTERNET



VARFUENTE: REVISTA SEMANA DEL 24 AL 31 DE OCTUBRE DE 2005 PAGINA 13



- 1) ¿Que variable se está estudiando en cada uno de los casos?
- 2) Nombre cada uno de los valores que toma la variable para cada uno de los casos.
- 3) ¿Encuentras alguna diferencia entre las variables que se están estudiando? ¿Cuál?
- 4) ¿Podríamos hacer una clasificación entre ambas variables? ¿Cuál?

Luego se pasaría a trabajar entre todo el grupo y ver las regularidades, juntarlas y llegar a la clasificación de las variables en cuantitativas y cualitativas.

### 5.7.2. Tarea 2. Variables en escala nominal y ordinal

Ahora bien se tomarán las cualitativas y se trabajará un poco sobre ellas, y ver una nueva clasificación:

GRAFICA 5.3.

RESULTADOS ENCUESTAS PARA PROXIMAS ELECCIONES



FUENTE: REVISTA SEMANA DEL 26 DE SEPTIEMBRE AL 3 DE OCTUBRE DE 2005 PAGINA 32

1. ¿Qué tipo de variable es la que se está estudiando en este caso?
2. ¿Puede darle un orden a los valores que toma la variable? Sin tener en cuenta la frecuencia, sino la importancia de la variable ¿Cuál orden le daría a excelente, a buena, a aceptable, a mala, y a muy mala?
3. ¿Si le pudiese dar valores numéricos a cada una de las opciones de respuesta como lo haría según su clasificación?

Luego de responder las preguntas se llega a una conclusión general y se presenta la siguiente gráfica:

GRAFICA 5.4.

RESULTADOS ENCUESTAS PARA PROXIMAS ELECCIONES



FUENTE: REVISTA SEMANA DEL 26 DE SEPTIEMBRE AL 3 DE OCTUBRE DE 2005 PAGINA 31

4. ¿Cuál es la variable de estudio en la gráfica 5.4?
5. ¿Podrían poseer un orden de importancia éstas? Sin tener en cuenta las frecuencias claro. ¿Por qué?
6. ¿Qué diferencia encuentra entre esta variable y la estudiada en el caso de la gráfica 5.3?

Tal y como su nombre lo indica las variables a las cuales se les puede dar un orden de acuerdo a algún criterio de ordenación son clasificadas dentro de la escala ordinal, en el caso del ejemplo en la tabla 5.3., los valores de las variables muy mala, mala, aceptable, bueno, excelente pueden tomar valores como 1, 2, 3, 4, 5, respectivamente, así pues la propiedad de orden que éstas poseen nos permite establecer relaciones entre ellas como mayor que, igual a, más rápido que, menos duro que, etc.

Aquellas variables que no es posible darles un orden se encuentran dentro de la escala nominal, éstas expresan cierta características y no poseen propiedades aritméticas, es decir

no tiene sentido sumarlas, multiplicarlas, etc., es por esta razón que son clasificadas dentro de la cualitativas.

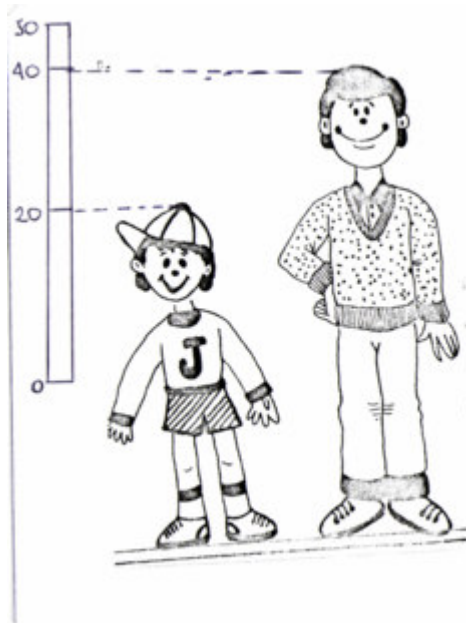
Ahora bien para el caso de las variables cuantitativas existe una primera clasificación que corresponde a las escalas que son utilizadas, estas son la escala de intervalo y la escala de razón.

### 5.7.3. Tarea 3. Variables en escala de intervalo y de razón

Se tiene la siguiente situación: Dos niños deciden determinar cuántos centímetros es uno más alto que el otro, pero solo tienen una regla de 50 cm., así marcan su estatura en la pared, y luego toman la regla desde un punto sobre la pared, a cierta distancia del piso que no tienen en cuenta. El de menor estatura es Juan y el otro es Pablo.

GRAFICA 5.5.

GRÁFICO ESTATURAS DE JUAN Y PABLO



- 1) ¿Cuál es la diferencia entre las alturas, de Juan y Pablo?
- 2) Si tuviesen un metro, y tomaran las medidas desde el suelo ¿la diferencia entre las medidas diferiría?

- 3) Supongamos que Juan mide 1.20 m. ¿Cuánto mide Pablo? ¿Cuál es la diferencia entre ambas medidas? ¿A que distancia fue puesta la regla del suelo en la gráfica?
- 4) ¿Es la estatura de uno el doble de la del otro?

De esta forma la escala de intervalo fue la que se utilizó al tomar la regla desde cualquier punto de la pared, es decir que hay un cero arbitrario, que fue tomado por uno de los niños sobre la pared, y para el caso de la escala de razón, existe un cero absoluto ya establecido, que en este caso es el suelo para hallar medida de estaturas. Como se puede ver en ambos casos la diferencia entre las medidas es la misma.

#### **5.7.4. Tarea 4. Practica 1**

Para trabajo en casa se pide que hagan un experimento cualesquiera en el cual pueda ser utilizada la escala de intervalo y la de razón.

##### *Explicación*

Como explicación se da que también existe otra clasificación de las variables cuantitativas, y es el caso discreto y continuo, en el primero de ellos los valores que se toman, poseen huecos, es decir una variable discreta deja de tomar valores de recta real entre dos enteros dados, en cambio la variable continua puede tomar todos los valores de un intervalo dado, así pues desde el punto de vista matemático los valores de una variable discreta podrán ser representados por intervalos cerrados o abiertos  $( )$  ó  $[ ]$ , pero en el caso de las variables discretas estas podrán ser representadas por medio de un conjunto  $\{ \}$ .

Un ejemplo podría presentarse para el caso de la variable discreta es el número de libros que salen para préstamo de la Biblioteca en un día, van a salir algunos de los valores del conjunto  $\{0, 1, 2, 3, \dots\}$ , pero nunca valores entre ellos, ya que un estudiante no podrá llevar  $\frac{1}{2}$  libro, y para el caso de una variable continua se tiene la cantidad de dinero en una cuenta en pesos, es posible tener \$15.819.75, sería posible afirmar que los clientes de cierta entidad bancaria tienen valores entre  $( \$5000, \$1000.000.000)$ . Hay valores entre ellos porque puede haber un cliente que tenga \$153.278.916.

### 5.7.5. Tarea 5. Variables continuas y discretas

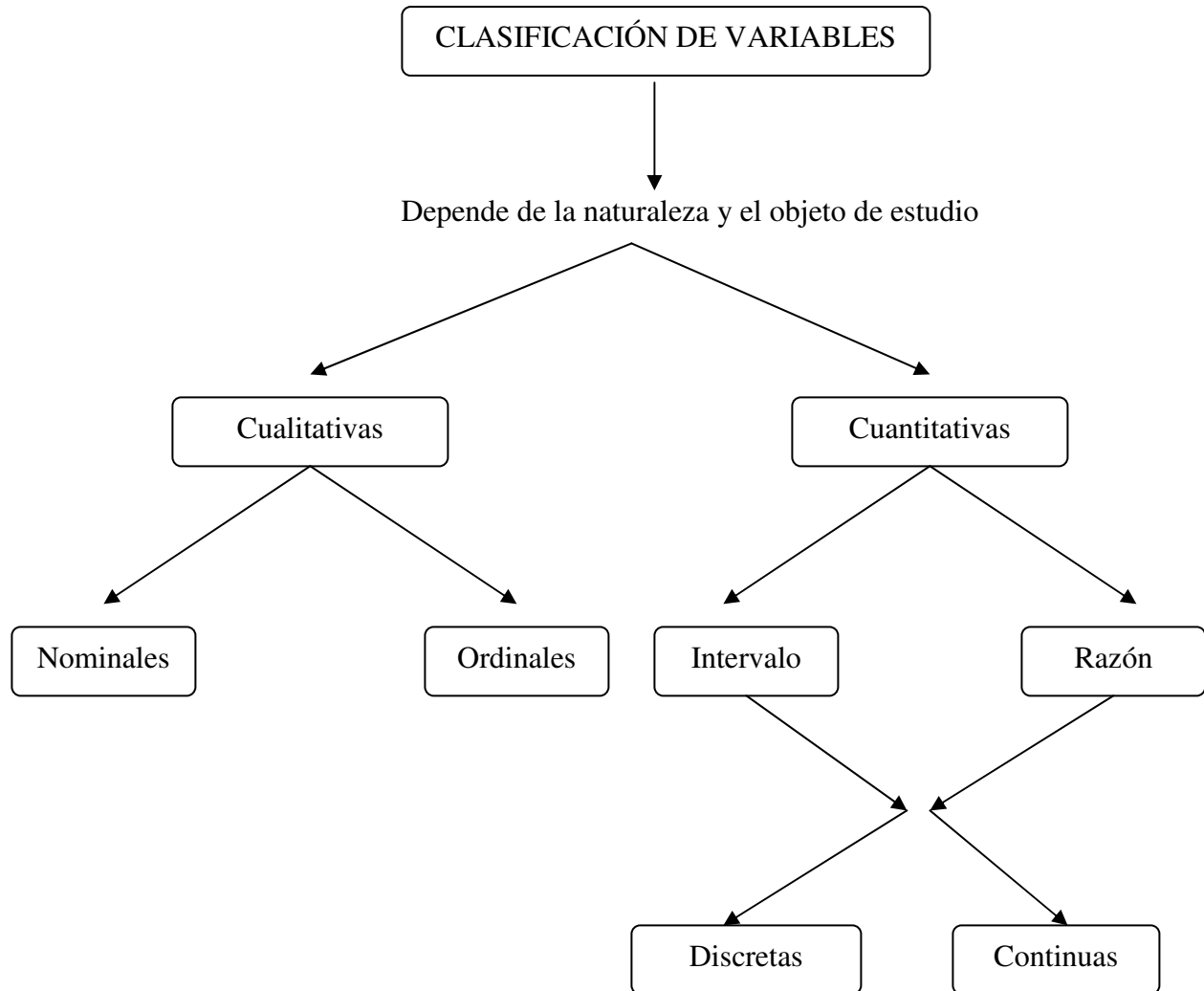
A continuación se les pide a los estudiantes clasificar las siguientes variables en continuas o discretas, según la definición y los ejemplos dados:

- 1) La estatura expresada en metros de sus compañeros
- 2) Resultados obtenidos al lanzar un dado
- 3) Número de estudiantes matriculados en un colegio
- 4) Edades expresadas en años de un colegio.
- 5) Número de accidentes automovilísticos que se producen en Bogotá en un día determinado.
- 6) Edad expresada en meses, días, minutos, etc., de los estudiantes de un salón de clase.
- 7) El tiempo de duración de una clase.
- 8) El peso de una persona.
- 9) Número de empanadas vendidas en la cooperativa del colegio.
- 10) Distancia que pueden recorrer diferentes personas en 5 minutos.
- 11) Número de profesores que laboran en un colegio

Así pues que los siguientes esquemas resultarán útiles para una primera clasificación de las observaciones que se han hecho sobre las variables:

TABLA 5.1.  
RESUMEN TIPOS DE VARIABLES

Nivel de medición Conceptualización matemática	Nominal	Ordinal	Intervalo	Razón
Noción de clasificación	Si	Si	Si	Si
Noción de orden	No	Si	Si	Si
Noción de densidad	No	No	Si	Si
Noción de distancia	No	No	Si	Si
Existencia de un cero absoluto	No	No	No	Si
Noción de continuidad	No	No	Si	Si



Otro aspecto que resulta importante nombrar, es que en un estudio se puede utilizar más de una variable.

**5.7.6. Tarea 6. Estudios univariados y multivariados**

Se tienen las siguientes gráficas:

TABLA 5.2.

**Atentados contra torres de energía de todas las empresas**

MES	2004	2005
Enero	3	4
Febrero	6	10
Marzo	8	11
Abril	10	11
Mayo	23	27
Junio	25	11
Julio	15	10
Agosto	11	12
Septiembre	10	36
Octubre	1	19
Noviembre	4	38
Diciembre	5	38
TOTAL	121	227

FUENTE: PERIÓDICO EL TIEMPO JUEVES 5 DE ENERO DE 2006

TABLA 5.3.

OPINIONES SOBRE LA IMAGEN DE URIBE PRESIDENTE DE COLOMBIA 2002 - 2006

**¿HA CAMBIADO SU IMAGEN SOBRE URIBE?**

EN CIFRAS	Septiembre 2005	Octubre 2005	Tendencia
Conocimiento	99%	100%	=
Favorabilidad	78%	72%	6% ↓
Favorabilidad frente a la reelección	68%	65%	3% ↓
Aprobación gestión	67%	62%	5% ↓
Apoyo	67%	62%	5% ↓
Percepción de cumplimiento	61%	58%	3% ↓
Intención de voto	56%	56%	=

En la mayoría de aspectos, en el último mes la imagen de Uribe bajó

FUENTE: REVISTA SEMANA DEL 21 DE NOVIEMBRE AL 28 DE NOVIEMBRE DE 2005 PAGINA 112

GRAFICA 5.6.

OPINIONES SOBRE LAS ELLECCIONES DE PRESIDENTE DE COLOMBIA 2006 - 2010



FUENTE: REVISTA SEMANA DEL 31 DE OCTUBRE AL 7 DE NOVIEMBRE DE 2005 PAGINA 40

1. ¿Que variables están involucradas en cada uno de los reportes de los estudios que se representan en las gráficas 5.6, 5.7, y 5.8?
2. ¿Qué diferencia encuentra entre ellas?
3. ¿Qué diferencia encuentra entre las variables de estos estudios y los que representan las gráficas 5.4 y 5.1?

Luego se puede concluir que los estudios que tienen dos variables son llamados vibariados y con más de dos variables multivariados.

### 5.7.7. Tarea 7. Recolección y organización de datos.

Para iniciar con el trabajo de distribución de frecuencias se pedirá a los estudiantes que hagan un estudio en el cual recojan los siguientes datos, entre 30 estudiantes de dos grados cualesquiera de su colegio, por medio de una encuesta, la cuál será representada mediante un cuadro como el que se presenta a continuación:



No	Nombre	Edad	Sexo	Estatura	Peso	Hermanos	Curso
1	...	...	...	...	...	...	...
2	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
30	...	...	...	...	...	...	...

### 5.7.8. Tarea 8. Frecuencias absolutas

Luego de que el estudiante haya recolectado esta información, se le pide que haga un recuento de los datos en común que haya obtenido, ya que para hacer un análisis más a fondo acerca de la información es necesario que halle las frecuencias de ocurrencia de los datos que recogió, es decir hallar el número de veces que se repite determinado valor. Y con estas realice 7 nuevas tablas en su casa, las cuales van a estar distribuidas de la siguiente manera:

Edad	Frecuencia
$X_1$	$f_1$
$X_2$	$f_2$
...	...
...	...
...	...
$X_{30}$	$f_{30}$

Sexo	Frecuencia
$X_1$	$f_1$
$X_2$	$f_2$
...	...
...	...
...	...
$X_{30}$	$f_{30}$

Hermanos	Frecuencia
$X_1$	$f_1$
$X_2$	$f_2$
...	...
...	...
...	...
$X_{30}$	$f_{30}$

Luego de llevar a cabo dicho trabajo, se harán las siguientes preguntas de acuerdo a lo observado.

- 1) ¿En qué cursos desarrolló su estudio? ¿De que curso hubo más cantidad de estudiantes encuestados?
- 2) ¿Entre qué edades se encuentran los estudiantes entrevistados? ¿Cuál fue la edad más frecuente? ¿Cuál la menos frecuente?
- 3) ¿Entrevistó más cantidad de mujeres o de hombres?
- 4) ¿Cuál es el número de hermanos que más se repiten en los estudiantes de los niveles que entrevistó?
- 5) ¿Existe alguna constante en los nombres de los estudiantes?
- 6) ¿Entre qué estaturas están los estudiantes entrevistados? ¿Cuál fue la estatura más frecuente? ¿Cuál fue la menos frecuente?

- 7) ¿Entre que pesos están los estudiantes entrevistados? ¿Cuál fue el peso más y menos frecuente?
- 8) Suma los valores de las frecuencias de cada una de las tablas. ¿Qué valor obtienes en cada una?

### *Explicación*

Esta frecuencia que fue hallada  $f_n$  es llamada frecuencia absoluta, y la suma de todas da como resultado el número total de datos obtenidos, en este caso, el número de estudiantes encuestados. Ahora bien veremos que otra clase de frecuencias hay para trabajar:

### **5.7.9. Tarea 9. Frecuencias relativas**

La idea es que el estudiante genere nuevas tablas en las cuales halle el porcentaje, respecto al total, de cada una de las frecuencias encontradas, este trabajo se puede hacer utilizando reglas de tres. Las nuevas tablas estarán diseñadas de la siguiente forma:

Estatura	Frecuencia	Porcentaje
$X_1$	$f_1$	$h_1$
$X_2$	$f_2$	$h_2$
...	...	...
...	...	...
...	...	...
$X_{30}$	$f_{30}$	$h_{30}$

<i>Peso</i>	<i>Frecuencia</i>	<i>Porcentaje</i>
$X_1$	$f_1$	$h_1$
$X_2$	$f_2$	$h_2$
...	...	...
...	...	...
...	...	...
$X_{30}$	$f_{30}$	$h_{30}$

Curso	Frecuencia	Porcentaje
$X_1$	$f_1$	$h_1$
$X_2$	$f_2$	$h_2$
...	...	...
...	...	...
...	...	...
$X_{30}$	$f_{30}$	$h_{30}$

- 1) Escriba lo que puede concluir al observar el porcentaje que representa  $h_2, h_9, h_{11}, h_{17}, h_{21}, h_{26}, h_{30}$  en cada uno de los aspectos que se refieren a la información de nombre, edad, sexo, estatura, peso, hermanos, curso de los 30 estudiantes encuestados.
- 2) ¿Qué valor obtiene al sumar los porcentajes de cada una de dichas tablas?

#### *Explicación 1*

Así pues los porcentajes obtenidos  $h_n$ , expresados ya sea como fracción o como decimal representan las frecuencias relativas de cierta observación, y al sumar las frecuencias relativas vamos a obtener siempre 100% si se está trabajando con porcentajes, 1 si es con decimales o fracciones.

Para terminar con el estudio de las frecuencias, se tiene el caso de las frecuencias acumuladas, es posible hallar la frecuencia absoluta acumulada, y la frecuencia relativa acumulada. Veremos el proceso para hallar dichas frecuencias acumuladas:

#### *Explicación 2*

Como se ha venido trabajando se tiene que  $f_n$  denota la frecuencia absoluta del dato  $X_n$  y  $h_n$  denota la frecuencia relativa de dicho valor, como su nombre lo indica la frecuencia acumulada se obtiene al ir sumando (acumulando) la frecuencia absoluta (relativa) del dato con todas las frecuencias absolutas anteriores a él.

### 5.7.10. Tarea 10. Frecuencia absoluta acumulada

Ahora bien a la frecuencia absoluta acumulada la denotaremos como  $F_n$  y a la frecuencia relativa acumulada como  $H_n$ . Así pues tendríamos para el trabajo que estamos realizando con los 30 estudiantes la frecuencia absoluta acumulada estaría representada por:

$$F_1 = f_1 \text{ (Porque no hay frecuencia absoluta anterior a } f_1)$$

$$F_2 = f_1 + f_2 = F_1 + f_2$$

$$F_3 = f_1 + f_2 + f_3 = F_2 + f_3$$

.

.

.

$$F_n =$$

### 5.7.11. Tarea 11. Frecuencia relativa acumulada

Según la regularidad de los valores en  $F_1, F_2, F_3, F_4$ , ¿Cómo podrías expresar  $F_n$ ?

Y la frecuencia relativa acumulada por:

$$H_1 = h_1 \text{ (Porque no hay frecuencia absoluta anterior a } f_1)$$

$$H_2 = h_1 + h_2 = H_1 + h_2$$

$$H_3 = h_1 + h_2 + h_3 = H_2 + h_3$$

$$H_4 = h_1 + h_2 + h_3 + h_4 = H_3 + h_4$$

.

.

.

$$H_n =$$

Para el caso de los estudiantes, tenemos que  $F_{30} = 30$  y  $H_{30} = 100\%$  ¿Por qué?

#### *Explicación 1*

Para terminar se le pedirá al estudiante que haga de nuevo las tablas que contengan las frecuencias absolutas, las relativas, las absolutas acumuladas, y las relativas acumuladas, teniendo como modelo así:

Estatura	Frecuencia Absoluta	Frecuencia Absoluta Acumulada	Frecuencia Relativa	Frecuencia Relativa Acumulada
$X_1$	$f_1$	$F_1$	$h_1$	$H_1$
$X_2$	$f_2$	$F_2$	$h_2$	$H_2$
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
$X_{30}$	$f_{30}$	$F_{30}$	$h_{30}$	$H_{30}$

Edad	Frecuencia Absoluta	Frecuencia Absoluta Acumulada	Frecuencia Relativa	Frecuencia Relativa Acumulada
$X_1$	$f_1$	$F_1$	$h_1$	$H_1$
$X_2$	$f_2$	$F_2$	$h_2$	$H_2$
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
$X_{30}$	$f_{30}$	$F_{30}$	$h_{30}$	$H_{30}$

Hermanos	Frecuencia Absoluta	Frecuencia Absoluta Acumulada	Frecuencia Relativa	Frecuencia Relativa Acumulada
$X_1$	$f_1$	$F_1$	$h_1$	$H_1$
$X_2$	$f_2$	$F_2$	$h_2$	$H_2$
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
$X_{30}$	$f_{30}$	$F_{30}$	$h_{30}$	$H_{30}$

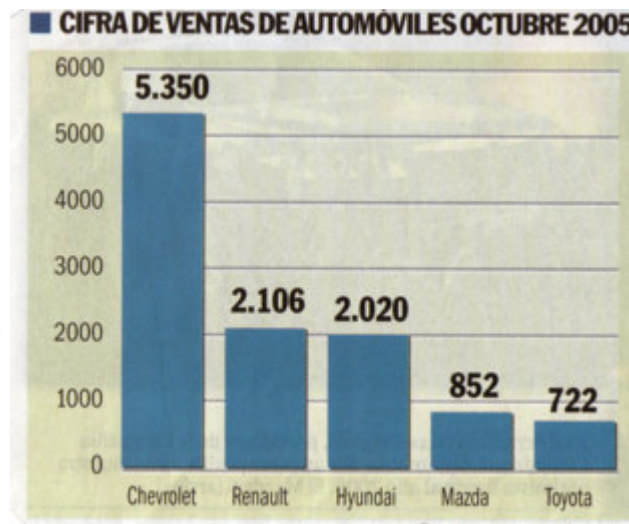
### Explicación 2

Continuando se les explica a los estudiantes que se han estudiado hasta ahora las tablas de frecuencia, pero que existe otra forma de presentar los datos de una situación problema que resulta más agradable para la vista, y en ocasiones más asequible para realizar análisis. Esta representación es la gráfica, algunos tipos de gráficas ya han sido vistas durante el desarrollo del tema de las variables, pero en esta parte se tratarán un poco más afondo.

#### 5.7.12. Tarea 12. Diagrama de barras vertical

Se inicia el trabajo con el diagrama de barras:

GRAFICA 5.7



FUENTE: REVISTA SEMANA 21 DE NOVIEMBRE DE 2005 PAGINA 131

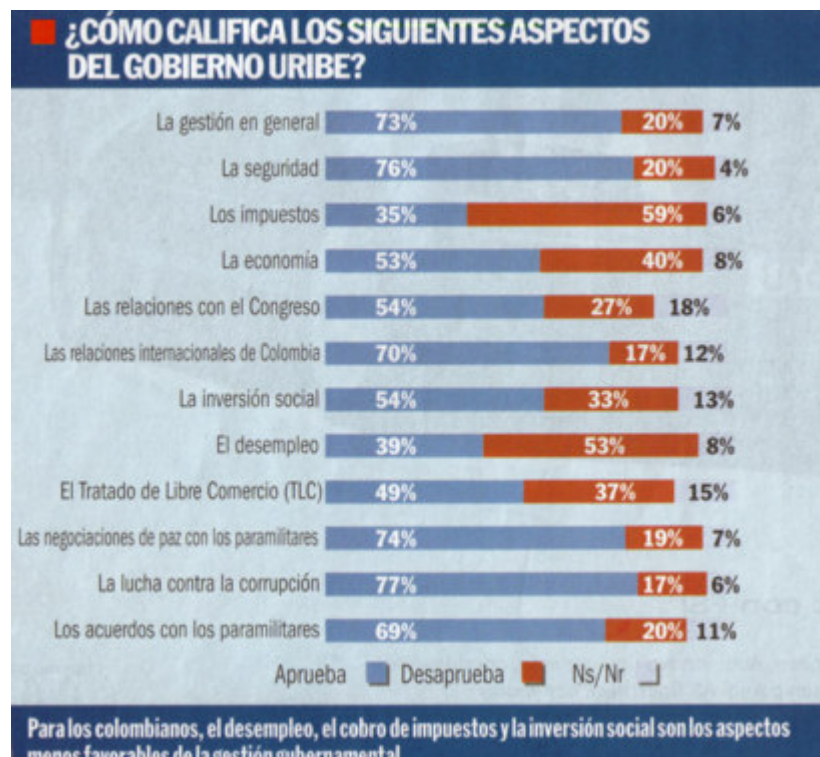
- 1) ¿Qué valores se toman en el eje x?
- 2) ¿Qué se está representando en el eje x?
- 3) ¿Qué se está representando en el eje y?
- 4) ¿Qué carro ha sido el más y el menos vendido en octubre de 2005?
- 5) Representa por medio de una tabla los datos dados en la gráfica 5.10.
- 6) En que caso resulta más fácil contestar la pregunta 4) ¿Por qué?

### 5.7.13. Tarea 13. Diagrama de barras horizontal

Se tiene la siguiente gráfica:

GRAFICA 5.8.

CALIFICACIÓN ACERCA DEL PRESIDENTE URIBE ELECTO 2005



FUENTE: REVISTA SEMANA DEL 24 AL 31 DE OCTUBRE DE 2005 PÁGINA 3



Respecto a ésta se propone el ejercicio de representar la información suministrada en la gráfica por medio de un diagrama de barras verticales sobre el plano cartesiano. Con esta tarea se podrá ver si el estudiante le da sentido a los porcentajes sobre el plano cartesiano.

#### **5.7.14. Tarea 14. Conclusión diagrama de barras**

Como se ha podido ver el diagrama de barras, resulta útil sobretodo para el caso de variables cualitativas, o cuantitativas pero no continuas, ya que estas están separadas sobre el eje x. Se culmina proponiendo un ejercicio, en el cual los estudiantes digan al menos tres características que posee un diagrama de barras. A continuación se presentan cinco de ellos:

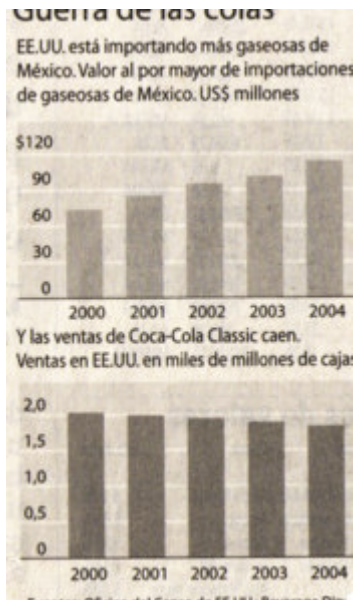
1. El diagrama de barras está construido sobre un par de ejes de coordenadas.
2. En el eje horizontal o x, se escriben los valores de la variable, y en el vertical se colocan sus respectivas frecuencias.
3. Las barras o rectángulos están separados uno del otro.
4. Todas las barras tienen la misma amplitud.
5. La altura de cada barra representa la frecuencia del dato correspondiente a ésta.

#### **5.7.15. Tarea 15. Practica 2**

Como trabajo se les deja hacer un informe escrito acerca de lo que está pasando con la Coca Cola en México y EE.UU. de acuerdo la gráfica 5.9:

GRAFICA 5.9

SITUACIÓN DE LA BEBIDA COCA COLA EN EE.UU. Y MÉXICO.



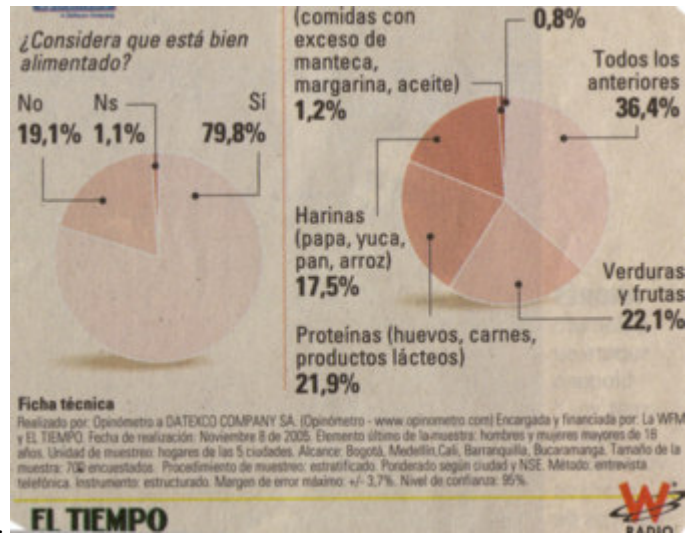
FUENTE: PERIÓDICO EL TIEMPO 11 DE ENERO DE 2006

### 5.7.16. Tarea 16. Diagrama circular

Continuando tenemos el diagrama circular, el cual normalmente se utiliza para representar frecuencias relativas observe el siguiente gráfico y conteste a las preguntas:

GRAFICA 5.10

ESTADO DE LA ALIMENTACIÓN DE LOS COLOMBIANOS



FUENTE: PERIÓDICO EL TIEMPO 11 DE ENERO DE 2006

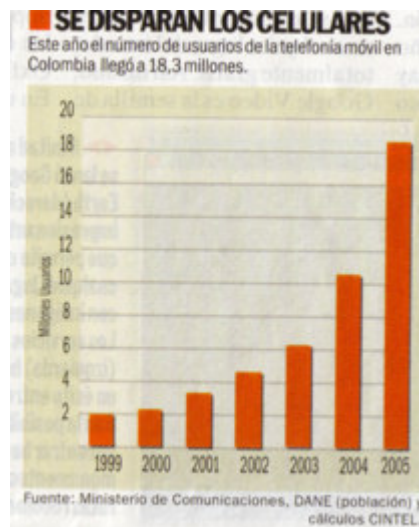
- 1) El número de personas que contestaron la encuesta fue 700. Halle las respectivas frecuencias absolutas correspondientes a cada una de las variables, y organízasalas por medio de una tabla.
- 2) ¿Considera usted que la población colombiana está bien alimentada? ¿Por qué?
- 3) Representa los datos mediante un diagrama de barras
- 4) ¿Qué podría concluir acerca de la encuesta?
- 5) ¿Cuál sería el cambio en los porcentajes, si el número de personas variara?

### 5.7.17. Tarea 17. Practica 3

Ahora bien se tienen los siguientes datos acerca del uso de celulares

GRAFICA 5.11

USO DE CELULARES



FUENTE: PERIÓDICO EL TIEMPO 2 DE ENERO DE 2006

Con base en esta información se le deja como trabajo al estudiante que construya un diagrama circular que la represente, y que analice la información acerca del aumento de usuarios en la telefonía móvil en Colombia.

Recordándole los siguientes aspectos:

- Una circunferencia tiene 360° de longitud
- Cada año representa una sección de la circunferencia ¿Cómo la hallaría?

### 5.7.18. Tarea 18. Pictograma

Otra representación es el pictograma, ésta ya no es muy utilizada hoy en día ¿Por qué? Pero se estudiarán algunas de sus propiedades. De acuerdo a la gráfica 5.15, conteste las siguientes preguntas:

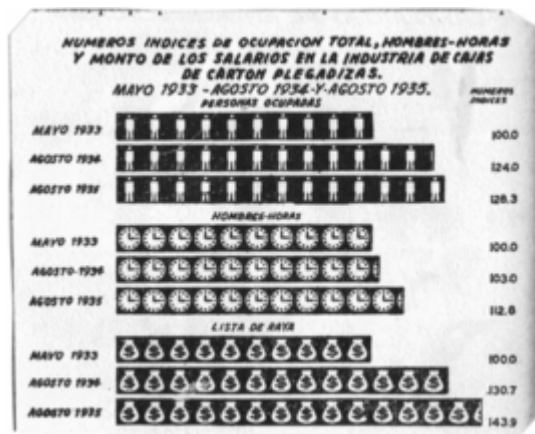
1. ¿En que año hubo una mayor cantidad de hombres trabajando?
2. ¿Qué características debe cumplir una gráfica para que sea un histograma? Nombre al menos tres de ellas de acuerdo con la gráfica.
3. Si no estuvieran los valores a la derecha de la gráfica ¿Sería posible determinar el valor exacto que están representando cada una de las gráficas?

Seis de las características que tiene un pictograma son:

1. Los símbolos se explican por si solos
2. Las cantidades mayores se indican por medio de un número mayor de símbolos, no por símbolos más grandes.
3. Los diagramas comparan cantidades aproximadas, no detalles minuciosos.
4. Los pictogramas se utilizan para hacer comparaciones, no afirmaciones aisladas.
5. Puede estar representado sobre un plano cartesiano.
6. Es posible representar frecuencias absolutas o relativas.

GRAFICA 5.12

PICTOGRAMA



FUENTE: ESTADÍSTICA GENERAL APLICADA (1992)

### 5.7.19. Tarea 19. Practica 4

Se deja como actividad, que a partir del diagrama de barras, y los datos de la tabla representen por medio de un pictograma cada uno de los grupos de los datos representados en éstos. Y además hagan una comparación, en cuanto a la exactitud y entendimiento por parte del lector de la información que arrojan cada una éstas gráficas dadas y las obtenidas. Y hacer un completo análisis acerca de la información dada.

GRAFICA 5.13

INFORMACIÓN DE EMPRESAS Y SU VALOR EN LA BOLSA



FUENTE: REVISTA SEMANA 26 DE DICIEMBRE DE 2005

TABLA 5.4.

OPINIÓN DE LOS CIUDADANOS ACERCA DE LOS POSTULADOS A LA CÁMARA

Germán Vargas Lleras	23,6%	Mario Uribe	3,5%
NS/NR	18,0%	Luis Alfredo Ramos	3,0%
Enrique Peñalosa	13,5%	Carlos Gaviria	2,5%
No votaría	8,5%	Juan Manuel Galán	2,0%
María Emma Mejía	6,5%	Claudia Rodríguez	1,5%
Gustavo Petro	5,5%	Armando Benedetti	1,0%
Ninguno de los anteriores	5,0%	Ricardo Lozada	0,5%
Juan Manuel Santos	5,0%	Alfonso Angarita	0,5%

FUENTE: REVISTA SEMANA 26 DE DICIEMBRE DE 2005

### 5.7.20. Tarea 20. Diagrama de tallo y hojas

Ahora bien, también se tiene el diagrama de tallo y hojas, para el cual se retomarán los 50 datos relacionados con las calificaciones obtenidas por los alumnos de un curso en el examen final de matemáticas:

TABLA 5.5.

CALIFICACIONES EXAMEN DE MATEMÁTICAS

70	68	90	73	91
75	75	82	81	83
60	76	81	66	73
54	70	71	44	94
40	42	78	81	90
90	70	67	57	86
83	86	69	51	88
76	71	55	54	90
70	54	53	58	74
68	52	62	64	76

Realice una tabla de dos columnas, en la primera de ellas ubique cada una de las decenas ordenadas de menor a mayor de arriba hacia abajo, y en la parte derecha, las unidades ordenadas de igual forma, pero de derecha a izquierda de tal manera que se formen con ambas columnas todos los valores dados en la tabla, no repita las decenas, ubique tantas unidades sea necesario en la parte derecha, separadas por un espacio.

Después que los estudiantes hayan realizado dicho trabajo, se procede a revisarlo:

TABLA 5.6.

CALIFICACIONES EXAMEN DE MATEMÁTICAS

Decenas	Unidades
4	0 2 4
5	1 2 3 4 4 4 5 7 8
6	0 2 4 6 7 8 8 9
7	0 0 0 0 1 1 3 3 4 5 5 6 6 6 8
8	1 1 1 2 3 3 6 6 8
9	0 0 0 0 1 4

Y así se obtiene, lo que llamamos un diagrama de tallo y hojas

GRAFICA 5.14

DIAGRAMA DE TALLO Y HOJA CALIFICACIONES EXAMEN DE MATEMÁTICAS

4	0 2 4
5	1 2 3 4 4 4 5 7 8
6	0 2 4 6 7 8 8 9
7	0 0 0 0 1 1 3 3 4 5 5 6 6 6 8
8	1 1 1 2 3 3 6 6 8
9	0 0 0 0 1 4

### Explicación

Recibe este nombre porque, las decenas representan el tallo, y las unidades las hojas. El diagrama doble de tallo y hoja, lo que hace es duplicar el número de posiciones del tronco, con la intención de buscar una mayor claridad en la presentación, de la siguiente forma.

GRAFICA 5.15

DIAGRAMA DOBLE DE TALLO Y HOJA CALIFICACIONES EXAMEN DE MATEMÁTICAS

4-	0 2 4
5-	1 2 3 4 4 4 4
5+	5 7 8
6-	0 2 4
6+	6 7 8 8 9
7-	0 0 0 0 1 1 3 3 4
7+	5 6 6 6 8
8-	1 1 1 2 3 3
8+	6 6 8
9-	0 0 0 0 1 4

### 5.7.21. Tarea 21. Análisis diagrama de tallo y hojas

Respecto a la tabla 5.4 y el diagrama 5.14 conteste las siguientes preguntas:

- 1) En donde resulta mejor leer la información en cuanto a aspectos como
  - Rapidez de análisis
  - Organización
  - Elaboración

Como se puede ver, los datos acerca de las 50 calificaciones son de gran magnitud, así pues que se pasará a desarrollar otro trabajo en el cual el estudiante se relacionará más con datos dados y la agrupación de éstos para así lograr un mejor análisis, el primer aspecto que se estudiará será su ubicación en una tabla por medio de grupos para así facilitar su análisis.

### 5.7.22. Tarea 22. Análisis de datos

Conteste las siguientes preguntas teniendo como información los datos de la tabla 5.4.

- 1) ¿Cuál fue la calificación más alta obtenida por los estudiantes?
- 2) ¿Cuál fue la calificación más baja obtenida por los estudiantes?
- 3) ¿Cuál es la diferencia entre las calificaciones obtenidas en los numerales 1) y 2)?



- 4) ¿Qué calificaciones tuvieron los cinco estudiantes de mayor puntaje?
- 5) ¿Cuál fue la calificación del décimo estudiante con mayor puntaje?
- 6) ¿Cuántos estudiantes tuvieron una calificación de 72 o mayor?
- 7) ¿Cuántos estudiantes obtuvieron una calificación menor de 85?
- 8) ¿Qué porcentaje de estudiantes obtuvieron una puntuación mayor de 64 y menor que 85?
- 9) ¿Qué calificaciones no obtuvieron los estudiantes?

Luego de que ellos respondan o intenten responder cada una de las preguntas se continúa con el trabajo de agrupar los datos en intervalos. Así pues que se continuará con dicho trabajo.

### 5.7.23. Tarea 23. Elaboración tabla. (Intervalos de clase)

Se tiene la siguiente tabla y el estudiante tendrá la tarea de completarla, basados en la tabla 5.4.

TABLA 5.7.

CALIFICACIONES EXAMEN DE MATEMÁTICAS ORGANIZADOS POR INTERVALOS DE CLASE

Calificaciones en grupos	Calificaciones contenidas en cada grupo en forma ordenada	Número de calificaciones que se encuentran en cada grupo
[40 – 45]	40, 42, 44	3
[46 – 51]		
[52 – 57]		
[58 – 63]		
[64 – 69]	64, 66, 67, 68, 68, 69	6
[70 – 75]		
[76 – 81]		
[82 – 87]	82, 83, 83, 86, 86	5
[88 – 93]		
[94 – 99]		
TOTAL		50

### *Explicación*

Los grupos, son llamados también intervalos de clase, y hay una forma para lograr que el ancho de todas las clases sean iguales, y además que este número sea acorde con el número de datos:

- 1) Lo primero que se hará será hallar la diferencia entre la calificación mayor y la menor:  $97 - 40 = 57$  y a este valor le añadimos 1 para obtener el número total de resultados, obteniendo 58. Este valor en forma general recibe el nombre de rango o recorrido, e indica el número de valores diferentes que puede tomar la variable, y en forma general, tenemos que:  $\text{Rango} = (\text{Mayor} - \text{Menor}) + 1$
- 2) Generalmente se utilizan de 10 a 20 intervalos de clase; dependiendo de la cantidad de datos, así pues que se escoge un número en este intervalo, en este caso se tomará 12.
- 3) Como siguiente paso se divide la cifra, obtenida en 1), es decir el rango, entre 15 (el número potencial de intervalos de clase) para hallar el ancho del intervalo:

$$57 \div 12 = 4.75$$

- 4) Si el número no es entero, se redondea este valor al número entero más cercano, en este caso sería entonces 5. Este valor se llama amplitud de los intervalos, y se denotará por la letra  $i$ . En general, el punto 3) y 4) en forma general será:

$$i = \frac{\text{Rango}}{\text{N}^\circ \text{ de Intervalos}}$$

- 5) Siguiendo se toma el dato más bajo, como el límite inferior del primer intervalo de clases y se agrega a este  $(5 - 1)$  en general  $(i - 1)$  en este caso 4 para obtener el puntaje máximo del primer intervalo de clase. Es decir nuestro primer intervalo será:

$$40 - 44$$

- 6) Por último el límite inferior de la clase siguiente, será el número entero consecutivo al máximo puntaje del intervalo de la clase anterior, es decir 45, y se le agrega de nuevo 4, así

$$45 - 49$$

- 7) Es posible tomar los extremos o no hacerlo, en el caso de no hacerlo se toma el intervalo como ( ) abierto, y de hacerlo sería tomar el intervalo cerrado [ ], caso en el cual se toman los extremos del intervalo tal y como se hizo en los casos anteriores.

#### 5.7.24. Tarea 24. Practica 5.

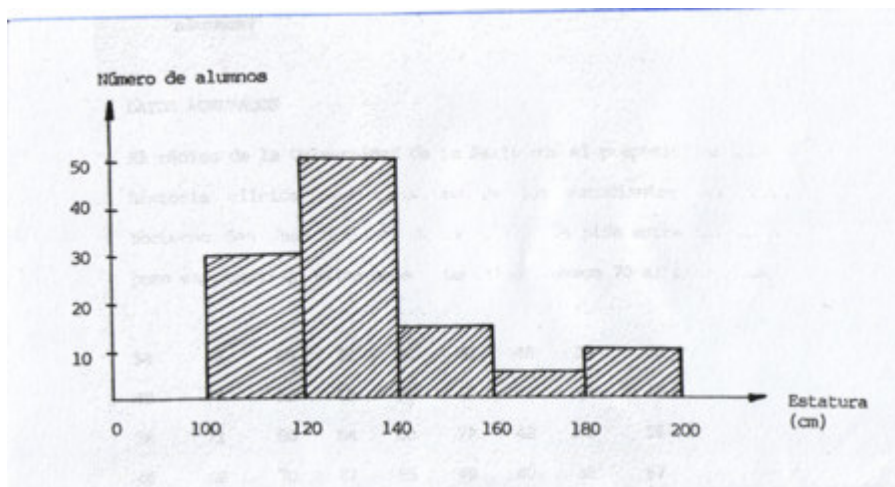
La tarea del estudiante será hacer la tabla correspondiente, con sus respectivas frecuencias, de cada uno de los 15 intervalos de clase que obtenga.

#### 5.7.25. Tarea 25. Histograma (Estatura vs. número de alumnos)

Ahora bien, la siguiente actividad consiste en que se tiene la siguiente gráfica:

GRAFICA 5.16

HISTOGRAMA



Y de acuerdo con ésta se plantean las siguientes preguntas:

- 1) ¿A qué se refiere el gráfico?
- 2) ¿Cuál es la variable que se está estudiando?
- 3) ¿Cuáles son los elementos de la muestra?
- 4) ¿Qué grupos de estatura escogieron? ¿Qué distancia abarca cada uno?

- 5) ¿Cuál es el valor de la estatura en cada uno de los extremos de los grupos formados?
- 6) ¿Cuántos alumnos se encuentran en cada uno de los grupos?
- 7) ¿En cuáles grupos se presenta el menor y el mayor número de alumnos?
- 8) Por medio de una tabla represente lo observado en las diversas preguntas

### *Explicación*

Se le explica al estudiante que es importante hallar un número que represente cada uno de los intervalos de clase, así pues que por conveniencia el que se utiliza es aquel que sea la mitad de cada intervalo, ¿Cómo hallarlo? Este número recibe el nombre marca de clase.

Para estos datos agrupados, se hallan de la misma manera todos los tipos de frecuencia. La idea es que los estudiantes elaboren una tabla que represente los datos dados en la gráfica 5.9., en la cuál se encuentren las frecuencias absolutas, relativas, absolutas acumuladas y relativas acumuladas.

Se les explica a los estudiantes que el histograma es un tipo especial de diagrama de barras, y este se encuentra representado en la gráfica 5.15, hay que tener en cuenta que con este se representan variables continuas, porque como se puede ver las barras van unidas en el eje x, es decir que no hay espacios en blanco, y es por esta razón que se representan los intervalos.

Ahora bien con el histograma, y la marca de clase de cada intervalo representado en éste, es posible hallar el polígono de frecuencias.

### **5.7.26. Tarea 26. Polígono de frecuencias**

Halle la marca de clase de cada uno de los intervalos representados en la gráfica 5.15, y únalos por segmentos, lo que obtienes se denomina polígono de frecuencias.

Veamos algunos de éstos:

GRAFICA 5.17

POLÍGONO DE FRECUENCIAS



FUENTE: PERIÓDICO EL TIEMPO 8 DE ENERO DE 2006

GRAFICA 5.18

POLÍGONO DE FRECUENCIAS

EXPORTACIONES DE BANANO



FUENTE: PERIÓDICO EL TIEMPO 5 DE ENERO DE 2006

### 5.7.27. Tarea 27. Análisis polígono de frecuencias

Con base en las gráficas 5.17 y 5.18 conteste las siguientes preguntas:

- 1) ¿Cuáles son las variables que se estudian en cada uno de los informes?
- 2) Represente por medio de una tabla los datos dados en los polígonos de frecuencia. Recuerde los intervalos de clase.
- 3) Rediseñe las gráficas con intervalos de seis meses en el caso de la tabla 5.18, y de 15 días para la gráfica 5.18.
- 4) Ahora reduce así sucesivamente los intervalos, hasta obtener curvas suavizadas.
- 5) ¿Puede ver alguna función representada en los datos?

### *Explicación*

Se les explica a los estudiantes, que si una curva de frecuencia tiene una forma parecida a la de alguna función matemática, entonces ésta se utilizará para describir todas las curvas parecidas a ésta y además se relacionará a una familia de distribuciones de frecuencias que tengan ciertos parámetros parecidos en la población que se esté estudiando, a través de las muestras.

## **5.8. Actividades del segundo tipo basadas en el uso de Minitools**

El programa “Minitools” fue creado con la meta de apoyar la enseñanza y el aprendizaje de la estadística. Sus herramientas y las secuencias educacionales asociadas están diseñadas para ayudar a los profesores en el proceso de enseñanza aprendizaje de la noción de distribución con los estudiantes. Es importante tener en cuenta el manejo del contexto social de los estudiantes y de los profesores, ya que los problemas que sean diseñados deben en lo posible estar guiados teniendo en cuenta éste.

La idea es llevar a cabo un análisis exploratorio de datos y desarrollar ciertas ideas de estadística por medio de las tres aplicaciones que proporciona Minitools. El foco en estadística en la educación media refleja un papel cada vez más importante del razonamiento estadístico, ya que varias revisiones (e.g., Garfield, 1988; Shaughnessy, 1992) revelan que la estadística recibe típicamente, en el mejor de los casos, una atención limitada en la escuela y que la instrucción se centra generalmente en aspectos de cómputo y procesales (e.g., calculando medias) estando así en la mayoría de los casos a expensas de la comprensión conceptual. Sin lugar a dudas, el análisis exploratorio de datos, en el cual los datos se manipulan y se indaga sobre ellos, para por medio de estos contestar a preguntas realistas o tomar decisiones es raramente el foco de la atención. La estadística en los libros de texto es dada para que los estudiantes experimenten con una actividad que implique el recordar ciertas operaciones con los números dados, mas no para llegar a hallar conjeturas y hacer un análisis más allá de lo que se ve.

Así pues a continuación se presenta una herramienta de innovación con la cual se estará dando un primer paso para llevar a cabo análisis de datos dentro del propio contexto en el que se este desarrollando el estudiante.

### **5.8.1. Manual introductorio de Minitools**

Se inicia dando algunas generalidades acerca del programa minitools, el cual fue el incentivo para el desarrollo del presente trabajo. La idea es dar una nueva herramienta al maestro para introducir en el aula de clase el tema de análisis exploratorio de datos y con este las herramientas para llevar a cabo un análisis de la distribución de frecuencias.


El CD contiene tres minitools estadísticos del applet. Fueron desarrollados por Paul Cobb, Koeno Gravemeijer y Janet Bowers en la universidad de Vanderbilt en los E.E.U.U., y programados por Michiel Doorman, Marco van Eck, Bastiaan Heeren, y Dick Wesseling en el instituto de Freudenthal en los Países Bajos, con el patrocinio del National Science Foundation de ESTADOS UNIDOS (Grant No. REC 9814898).

Al abrir el CD se encuentran dieciocho carpetas, las cuales están relacionadas con el sistema operativo del programa, ya para trabajar en las aplicaciones se abren los iconos de Internet Explorer. Para ejecutar el programa es necesario tener java, ya que las gráficas que son las herramientas principales para el uso de dicho software, solo pueden ser leídas con ese sistema.

Es posible encontrar tres niveles, minitool 1, minitool 2 y minitool 3, a continuación se llevará a cabo una explicación específica de cada uno de éstos

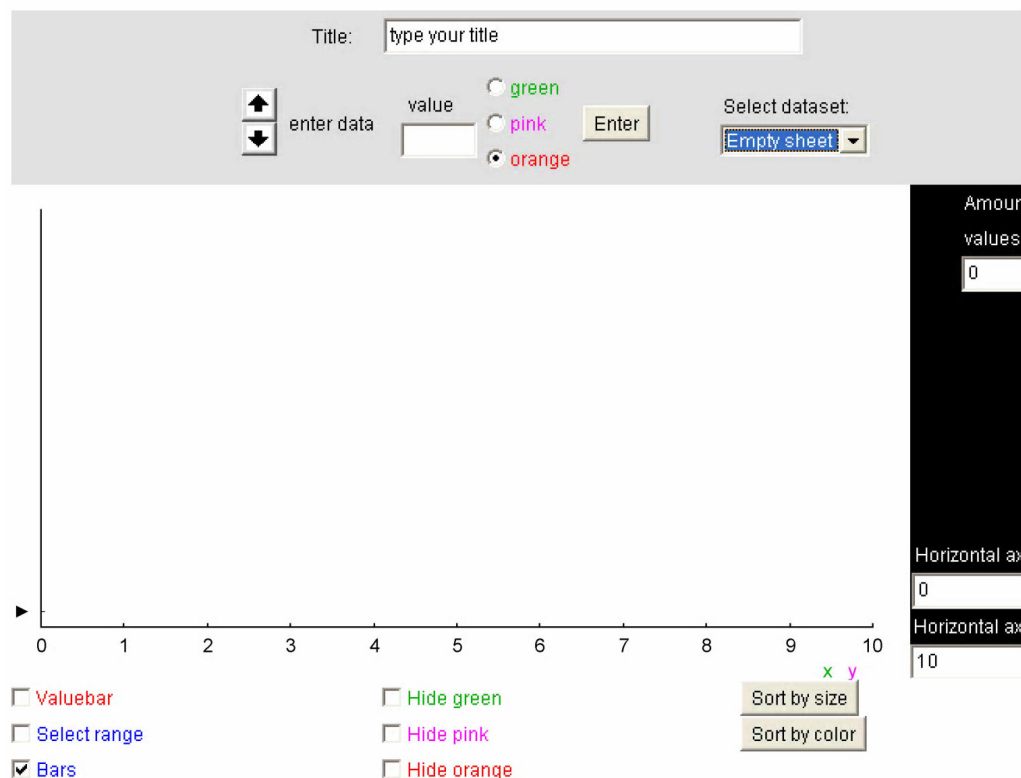
#### *Minitool 1*

Con esta primera herramienta se inicia con el trabajo de agrupación de datos, entonces se encuentra en primer lugar de arriba hacia abajo una ventana, que lleva el nombre de

“*title*”,  en la cual esta la opción de escribir un título para la gráfica que se va a insertar. (Gráfica 5.19.)

El siguiente paso a seguir es escoger un conjunto de datos, los cuales ya están programados, para este primer trabajo existen ocho opciones, la cuales están en ingles, y tienen que ver con distancia de frenado, libros, migraña, computadores, baterías, puntuaciones de cierta prueba, puntuaciones de un test, y un tanque de gasolina, y los colores con que aparecen las barras tienen que ver con diferentes poblaciones estudiadas. Así pues que la primera tarea del docente es diseñar situaciones problema que tengan que ver con los temas propuestos.

GRÁFICA 5.19.  
VISTA PRELIMINAR MINITool 1





Para continuar se encuentra la opción “*enter data*” (entrar datos), la cual se usa para ver los

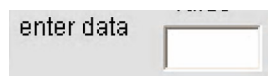


valores que toma la variable, dando clic en las flechas, el cursor se va desplazando de arriba hacia abajo o al contrario, mostrando así que valor esta tomando la variable en cada uno de los puntos, dentro de la caja que se encuentra en la parte derecha de dicha opción la



cual lleva en la parte superior el nombre de “*value*” (valor) va apareciendo el valor representado en eje horizontal, además es posible aumentar o suprimir cuantos valores se desee, escribiéndolos, o borrándolos de la caja, éstos aparecerán automáticamente dentro de la gráfica o desaparecerán.

Luego se encuentran tres colores, “*green*” (verde)  *green*, “*pink*” (rosado)  *pink*, “*orange*” (anaranjado)  *orange*, los cuales pueden ser usados en las barras que indican las frecuencias de las variables. Usando la opción “*enter data*” (entrar dato)



el cursor se ubica sobre cualquier barra, y así es posible cambiar de

color del valor que uno desee, escogiendo éste y dando clic sobre el icono “*enter*” , dependiendo el problema.

En la parte derecha de la ventana se encuentran tres datos acerca de la gráfica, el primero de arriba hacia abajo es “*Amount values*” (cantidad de valores), como su nombre lo indica, contabiliza el numero de valores que está tomando la variable, es decir controla lo que está mostrando el eje vertical dentro de la gráfica, éste valor depende del esquema no es posible hacerle cambios manualmente, varía dependiendo los valores que se adhieran o eliminen con la función vista anteriormente. Luego se encuentran “*horizontal axis min*” (valor mínimo en el eje horizontal), “*horizontal axis max*” (valor máximo en el eje horizontal), éstos determinan el valor mínimo y máximo que toma la variable y de esta forma da un rango para dichos valores. Al contrario de la función anterior ésta última si es necesario

controlarla manualmente, es decir de los valores que se den van a depender las variaciones de la gráfica.

Hasta ahora se han visto las opciones para hacer cambios dentro de la gráfica respecto a la forma, y los datos. Ahora bien para continuar en la parte inferior se encuentran herramientas para hacer algunos análisis de los datos.

En primer lugar se encuentra “*value bar*” (barra de valor)  Valuebar , dando clic en éste icono va a aparecer dentro de la gráfica una barra de color rojo que se puede desplazar de derecha a izquierda o viceversa, con un valor en la parte superior, el cual va a depender del valor que está tomando la variable en el punto donde se encuentre la barra. Esta opción va a ser útil para ver rangos y frecuencias absolutas acumuladas.

Luego está “*select range*” (selección de un rango)  Select range . Con esta opción se generan dos barras, las cuales pueden ser desplazadas sobre el eje horizontal tomándolas de la parte inferior para escoger un rango de valores, éste aparece en la parte superior de las barras, sobre cada una de ellas se encuentra el valor inicial y el valor final, y en la mitad de ambas se encuentra la cantidad de valores que se encuentran dentro de ese rango, además las barras que estén allí toman un color diferente al de las demás. Esta herramienta es útil para el trabajo con clases de frecuencias, y así con frecuencias relativas, y relativas acumuladas. (Gráfica 6.20.)

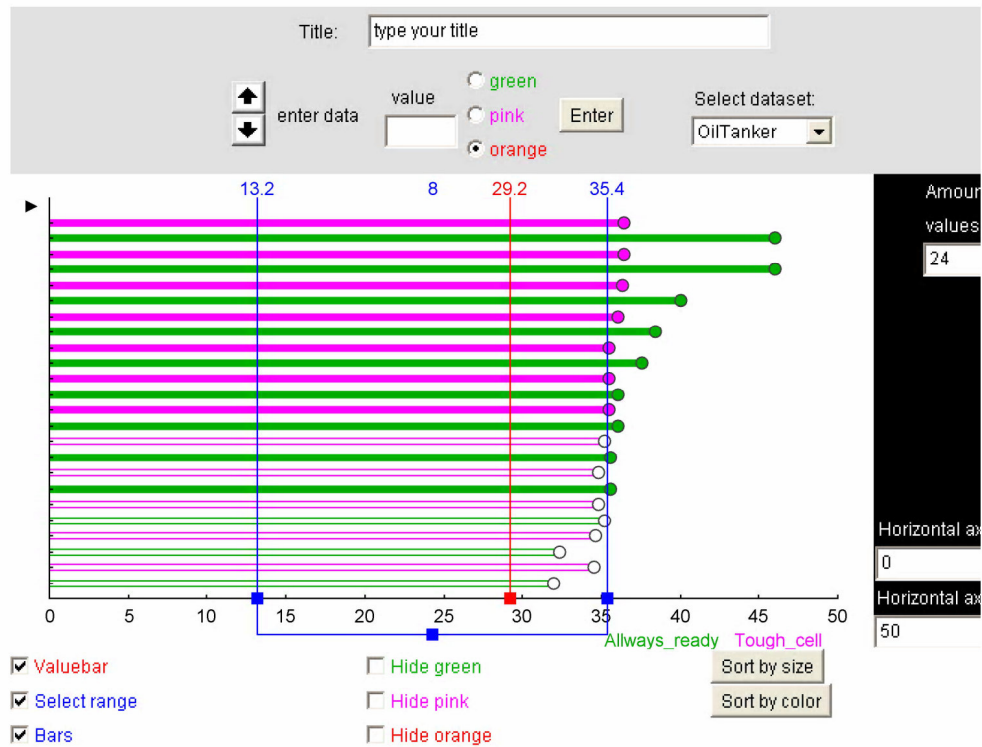
En el siguiente icono se encuentra “*bars*”  Bars (barras), opción con la cual es posible eliminar las barras y dejar solamente puntos, o al contrario, esto va a depender del tipo de análisis que se quiera el estudiante lleve a cabo. (Gráfica 5.21.)

Continuando se encuentran “*hide green*” (ocultar verde)  Hide green , “*hide pink*” (ocultar rosado)  Hide pink , y “*hide orange*” (ocultar anaranjado)  Hide orange , como su nombre

lo indica lo que va a hacer esta función será ocultar los valores de los colores seleccionados, dándole paso a un análisis univariado.

GRÁFICA 5.20.

REPRESENTACIÓN DE FRECUENCIAS CON BARRAS DE VALORES Y SELECCIÓN DE RANGO



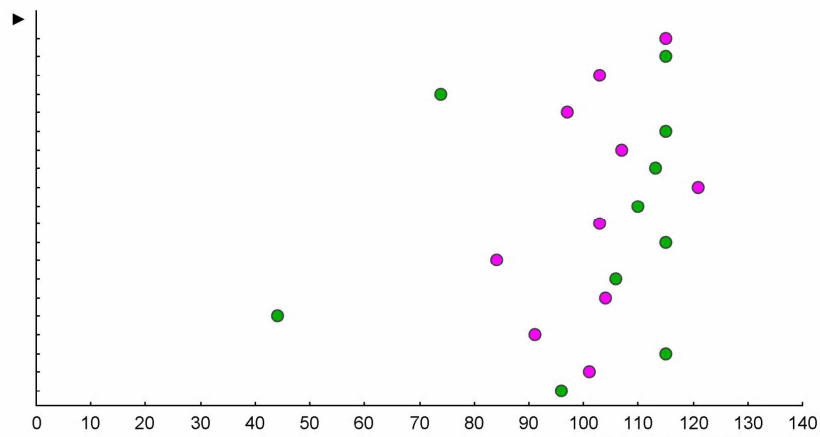
Para terminar los iconos “*sort by size*” (clasificar por tamaño) y “*sort by color*” (clasificar por color), van a organizar por tamaño de menor a mayor de arriba hacia abajo, o por colores las barras de datos. Lo que va a permitir hallar otra clase de rangos respecto a la cantidad de variables.

### Minitool 2

Los conjuntos de datos difieren, se tienen las variables, ambulancia, deseos, baterías, colesterol, muertes, fertilizantes, ritmo cardiaco 5, ritmo cardiaco 10, Madison, migraña, dinero, Nashville 1, Nashville 2, periódico, tanque de gasolina, fénix, reciclaje, respuesta,

San Francisco, Sat porcentaje, Sat puntuación, comida rápida, velocidad de engaño, ciudadanía. Y los colores que aparecen en los datos representan la variable  $x$  y la variable  $y$ . A diferencia del minitool 1, en este caso ya no aparecen barras, solo puntos, ya que se encuentran frecuencias mayores para cada uno de los valores de la variable. (Gráfica 5.22.)

GRÁFICA 5.21.  
REPRESENTACIÓN DE FRECUENCIAS CON PUNTOS



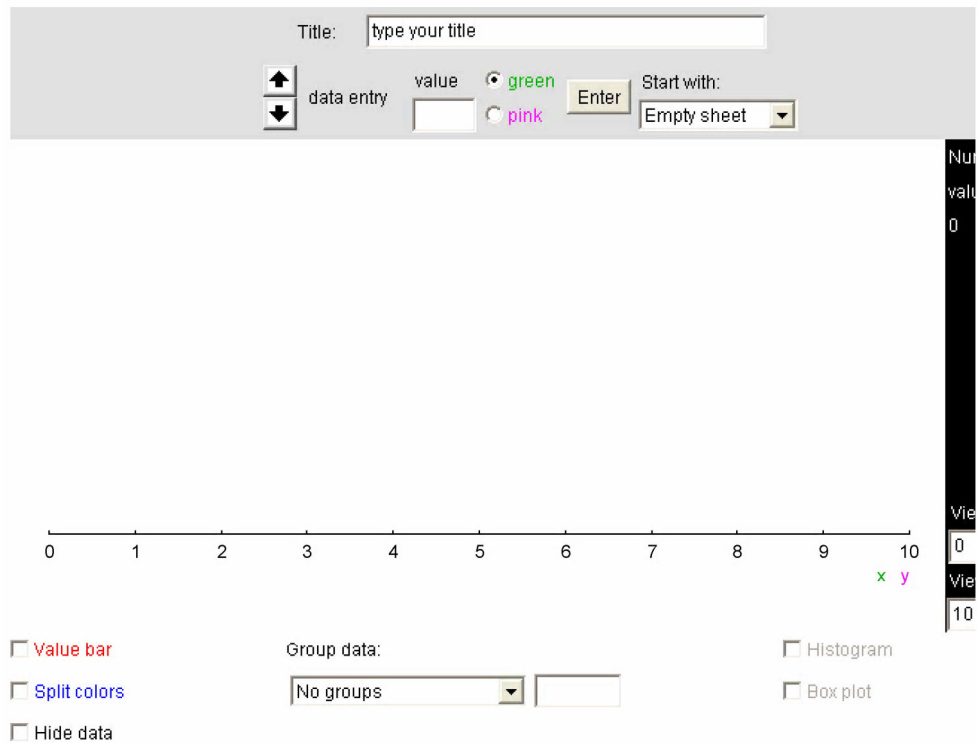
Al igual que en la parte uno, hay un espacio para darle un título a la gráfica, y continuando está el icono de **“data entry”** (ingreso de datos), haciendo clic sobre las flechas se van resaltando los puntos y así en el espacio para **“value”** (valores) van apareciendo los diversos valores que toma la variable en la posición de dicho punto.

Continuando hacia la derecha se encuentran los dos colores **“pink”** (rosado), y **“green”** (verde), los cuales se van marcando a medida que se van recorriendo los valores con la función anterior, o seleccionando el punto con las flechas, se toma un color y se da enter, inmediatamente el color del punto cambia por el seleccionado.

En la parte derecha se encuentra **“number of values”** (número de valores), donde aparece el número de valores con que se está trabajando. Continuando hacia abajo está **“view min.”**

(Ver mínimo) y “*view max.*” (Ver máximo), donde se define el rango de los valores que pueden tomar las variables.

GRÁFICA 5.22.  
VISTA PRELIMINAR MINITool 2




En la parte inferior se encuentra en primer lugar de arriba hacia abajo y de izquierda a derecha la opción “*value bar*”, al igual que en la parte uno, es una barra que tomándola de la parte inferior, es posible moverla sobre el eje horizontal de valores, y en la parte superior aparece el número en el cual se encuentra, permitiendo así hallar frecuencias absolutas.


Continua con “*split colors*” (partir colores)  *Split colors* dando clic sobre esta función, se van a dividir los datos de las dos variables, una en la parte superior y otra en la parte inferior, obteniendo así dos distribuciones univariadas. Con “*hide data*” (ocultar datos)  *Hide data*, los puntos de la gráfica se ocultarían, de esta forma solo se trabajaría con un grupo de los valores dados.

Desde este punto, se encuentran las diferencias con minitool 1, ya que se tiene la función

Group data:

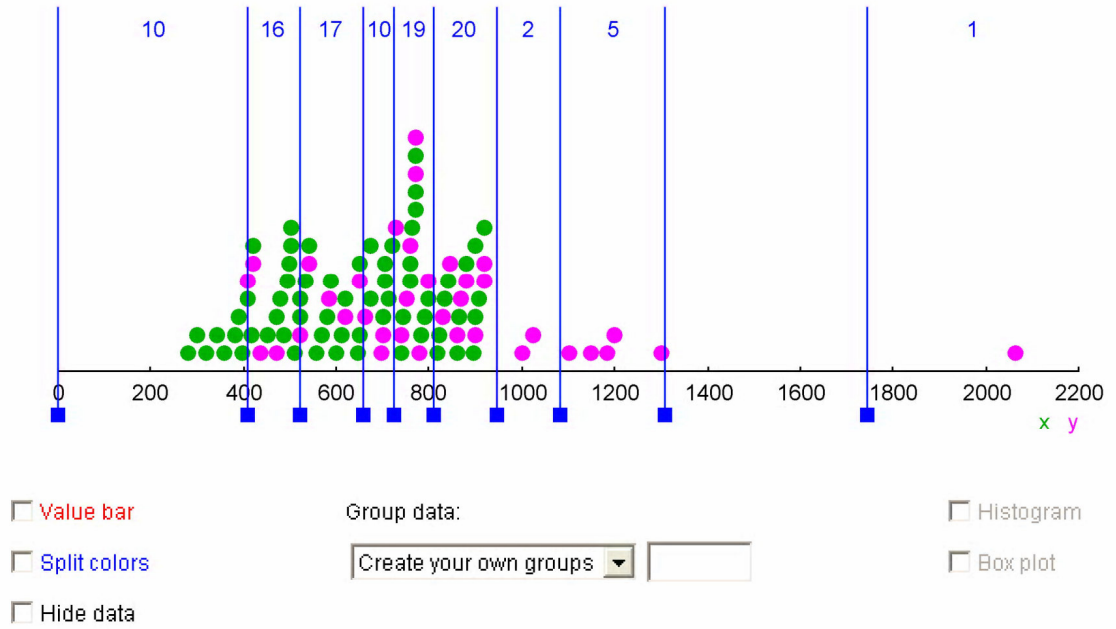
**“group data”** (grupos de datos) , como su nombre lo indica se usa para hacer diversos conjuntos de datos, al lado derecho de ésta se encuentra un número que indica la cantidad de conjuntos en que están divididos los datos.

Dentro de la anterior función es posible encontrar unas subfunciones que son **“create your own groups”** (crear sus propios grupos), dando clic sobre ésta aparece sobre la gráfica, al inicio del eje horizontal, una barra de color azul, si se hace clic sobre ella en la parte inferior y se mueve hacia la derecha se va a generar otra barra azul, y de esta forma se obtendrán dos grupos, dicha nueva barra se puede mover cuanto se quiera hacia la derecha, dependiendo del tamaño que se quiera la clase, y de así para formar tantos grupos como se desee a partir de la primera barra azul es posible hacerlo. En la parte superior de las barras se encontrará el número de elementos o frecuencia de datos de cada uno de los grupos. (Gráfica 5.23.)

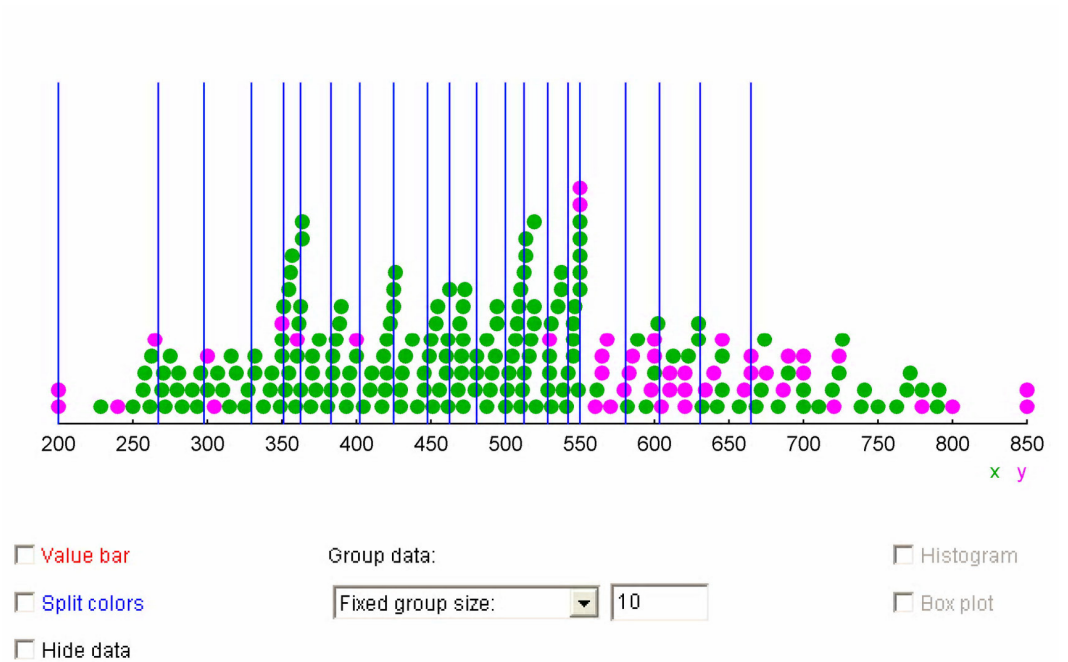
Sigue con **“fixed group size”** (fijar tamaño del grupo), lo que se hace es tomar una cantidad de puntos, frecuencia, valor que se escribe en el recuadro  y así se podría determinar cuantos grupos es posible formar con este número determinado de datos, ya que los conjuntos quedarían divididos mediante rectas verticales de color azul. (Gráfica 5.24.)

Continuando con **“fixed interval width”** (fijar ancho del intervalo), es posible tomar un ancho para el intervalo, dependiendo del número de datos, el cual se escribe en el recuadro, y automáticamente se generan barras de color azul que dividirán la gráfica dependiendo del rango dado, en los grupos correspondientes. Al darle clic sobre ésta se activará la función de **“histogram”**, (histograma)  Histogram la cual había estado desactiva hasta el momento, esto se debe a que el ancho de los intervalos va a ser el mismo para todos, y éste es un requisito fundamental para la elaboración de éste tipo de gráfica.

GRÁFICA 5.23.  
DATOS DIVIDIDOS EN NUEVE GRUPOS



GRÁFICA 5.24.  
DATOS DIVIDIDOS CON DIEZ FRECUENCIAS CADA GRUPO

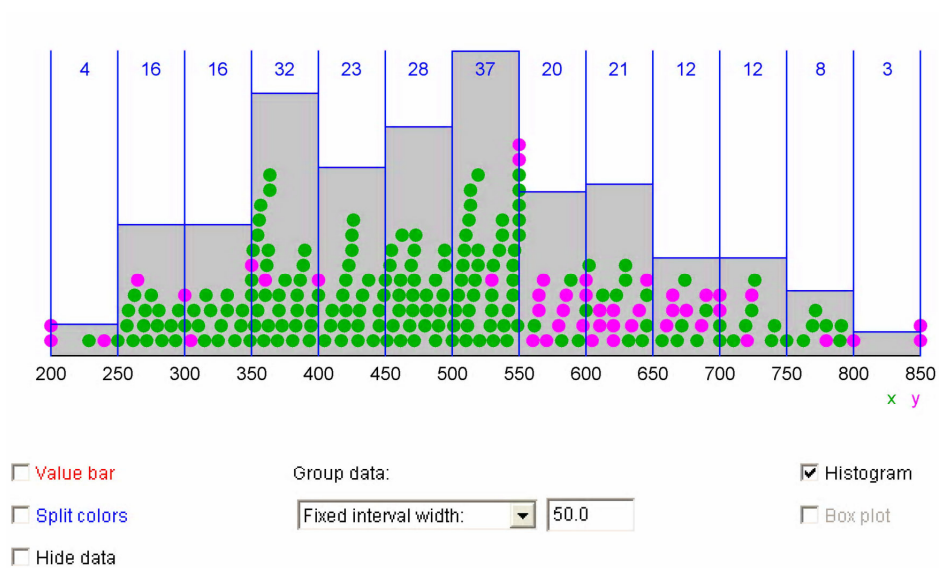


Si se da clic sobre ésta se generaría un histograma dentro del a gráfica sobre los puntos, dejando estos visibles, tendrá de alto tanto como puntos se encuentren dentro del intervalo. (Gráfica 5.25.)

Luego se trabaja con “*two equal groups*” (dos grupos iguales), función que dividirá la gráfica en dos grupos de frecuencias iguales, con lo cual se pueden hallar ciertas medidas de dispersión.

GRÁFICA 5.25.

HISTOGRAMA CON FRECUENCIAS DE CLASE DE 50 DE ANCHO CADA UNA.



Para terminar se tiene “*four equal groups*” (cuatro grupos iguales), con lo que se formarán cuatro grupos iguales y se generará automáticamente lo que se llaman cuartiles y así se activará “*box plot*” (caja de puntos). Ejercicio con el cual el estudiante podrá analizar y entender los términos cuartil y como se genera una caja de puntos. (Gráfica 5.26.)

### Minitool 3

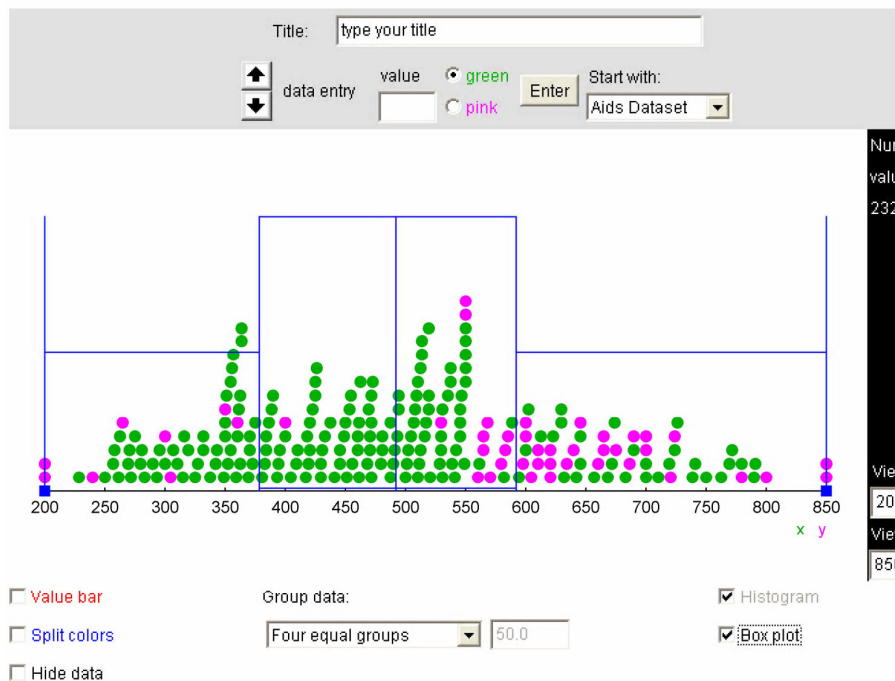
En este momento se inicia con el trabajo de análisis de variables un poco mas avanzado, porque se trabaja con dos variables al mismo tiempo, sin la diferencia de colores, ya que una estará ubicada en el eje horizontal y la otra en el eje vertical, ya se estarán relacionando una con la otra. (Gráfica 5.27.)



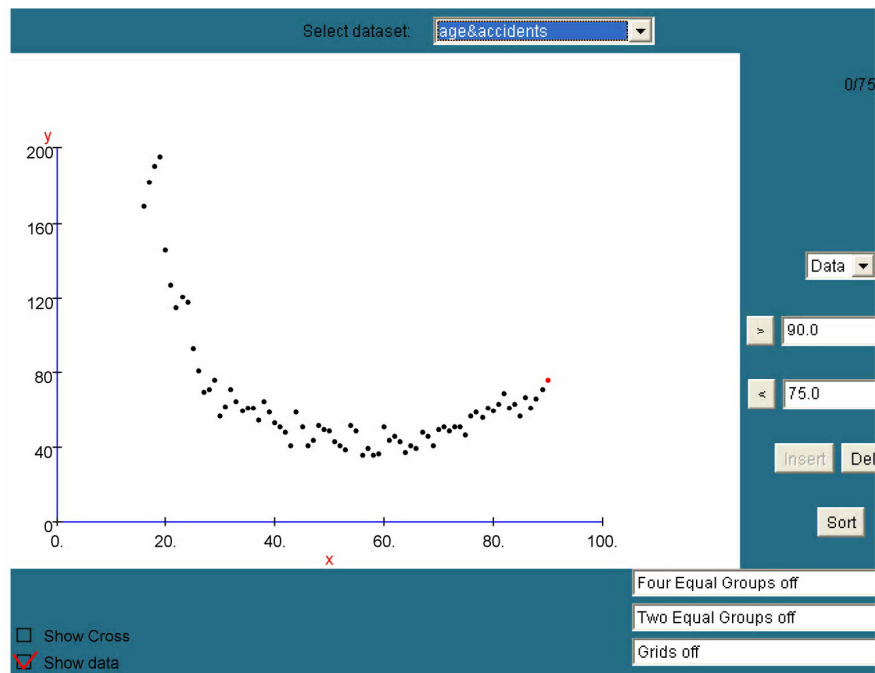
En primer lugar se encuentra la función “*select data set*” (selección del conjunto de datos), con lo que se escogerán los datos con los cuales se quiere trabajar, entre los que se encuentran, la relación entre edad y accidentes, edad y reacción, inquietudes, cepillado y placa bacteriana tipo A, cepillado y placa bacteriana tipo B, hueso de mineral tipo A, hueso de mineral tipo B, calorías y energía, oxígeno de 1958 a 1997, oxígeno cada año, oxígeno por tiempo, crimen y cantidad, crimen e ingreso, arrestos y edad, educación y salario M veinte, educación y salario M 50, educación y salario W 10, educación y salario W 20, educación y salario W 50, energía, fertilizante, horas durmiendo y primer test de puntos, horas durmiendo y segundo test de puntos, horas durmiendo y tercer test de puntos, velocidad de lectura grupo uno, velocidad de lectura grupo dos, trafico en 1992, trafico en 2000, neumáticos en la ciudad de San Francisco, neumáticos en la ciudad de Phoenix, nuevo.

GRÁFICA 5.26.

CAJA DE PUNTOS



GRÁFICA 5.27.  
VISTA PRELIMINAR MINITool 3



Como vemos en este último minitool se presenta una gran variedad de clases de datos, es necesario aclarar que las gráficas que aparecen ya son nubes de puntos. En la parte derecha de la gráfica se encuentra la cantidad de datos, continuando hacia abajo es posible ver un recuadro con la palabra “*data*” (dato), desplegándose de este otras dos opciones “*min.*” (Mínimo), y “*max.*” (Máximo). Continuando hacia abajo se encuentran dos recuadros blancos cada uno de ellos con una flecha a su lado izquierdo y al lado derecho un aviso en el primer recuadro de “*hor.*” (Horizontal) y en el segundo de “*ver.*” (Vertical). Así pues escogiendo la opción “*data*”, aparecerán en estos últimos recuadros valores correspondientes a la gráfica en el eje horizontal y su correspondiente valor en el vertical, y con las flechas se podrá desplazar sobre dichos valores, dando así a cada valor del eje horizontal uno en el eje vertical, en la gráfica cambiarán de color los puntos cuando estos vayan estando dentro del recuadro. Dando clic sobre “*min.*” o “*max.*” en los recuadros aparecerá respectivamente el valor mínimo o máximo de las variables representadas en el eje vertical y horizontal.

Luego se encuentra el recuadro llamado ***“delete”*** (borrar), dando clic sobre este se borrará el punto sobre el cual se encuentra el cursor, es decir el que esta seleccionado con la opción vista anteriormente. En el caso de tener los valores mínimo o máximo, no se activaría esta función.

Siguiendo hacia abajo, está un cuadro con el nombre ***“four equal groups off”*** (cuatro grupos iguales de), del cual se despliegan las opciones ***“4 x equal groups”***, ***“5 x equal groups”***, ***“6 x equal groups”***, ***“7 x equal groups”***, ***“8 x equal groups”***, ***“9 x equal groups”***, ***“10 x equal groups”***, con las cuales se dividirán los datos, respecto al eje horizontal en tantos grupos como indican los números, y lo que indica con el número cuatro, es que el eje vertical quedará dividido dentro de éste en 4 grupos de datos. (Gráfica 5.28.)

Luego es posible encontrar otro recuadro con el nombre de ***“two equal groups off”*** (dos grupos iguales de), del cual se despliegan las opciones ***“4 x equal groups”***, ***“5 x equal groups”***, ***“6 x equal groups”***, ***“7 x equal groups”***, ***“8 x equal groups”***, ***“9 x equal groups”***, ***“10 x equal groups”***, con las cuales se dividirán los datos, respecto al eje horizontal en tantos grupos como indican los números, y el eje vertical en dos grupos iguales de datos. (Gráfica 5.29).

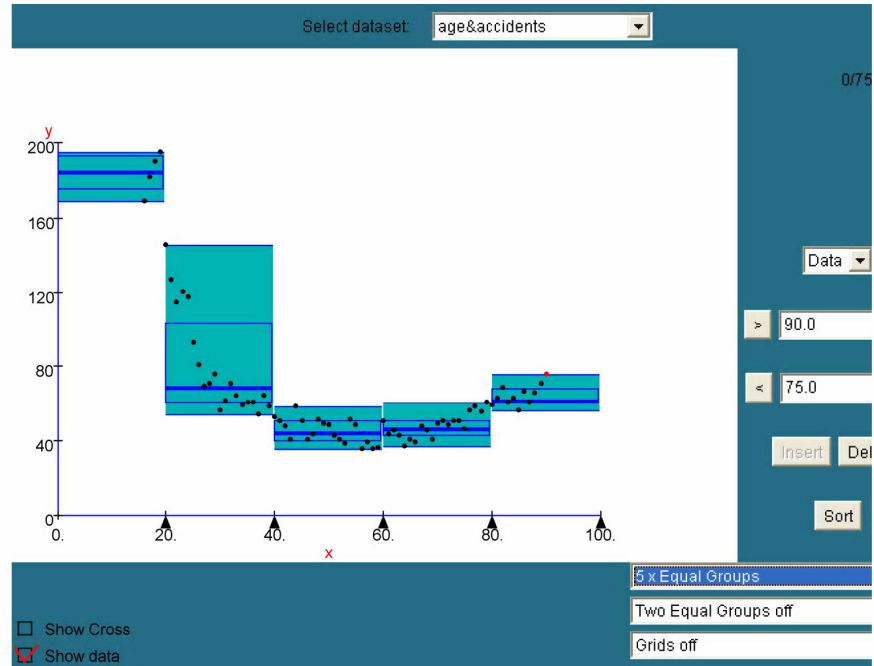
En el la última parte de la derecha hacia abajo se encuentra ***“grids off”*** (cuadrícula de), de donde se despliegan los valores ***“3x3”***, ***“4x4”***, ***“5x5”***, ***“6x6”***, ***“7x7”***, ***“8x8”***, ***“9x9”***, ***“10x10”***, opción con la cual se generará una tabla de las dimensiones dadas con los números correspondientes, la cual ocupa todo el espacio del conjunto de datos, además dentro de cada cuadro aparece el número de datos que se encuentran dentro de éste. (Gráfica 5.30.)

Para terminar en la parte inferior se encuentran dos opciones ***“show data”*** (mostrar datos), ***“show cross”*** (mostrar cruz), con las cuales puede llegar a aparecer o desaparecer el conjunto de datos, o un eje de coordenadas en el conjunto de datos, el cual tiene como valores los ya dados. Este eje se puede desplazar a partir de su origen hacia arriba, abajo,

derecha o izquierda dependiendo de los valores que se deseen hallar con esta. (Gráfica 5.31.)

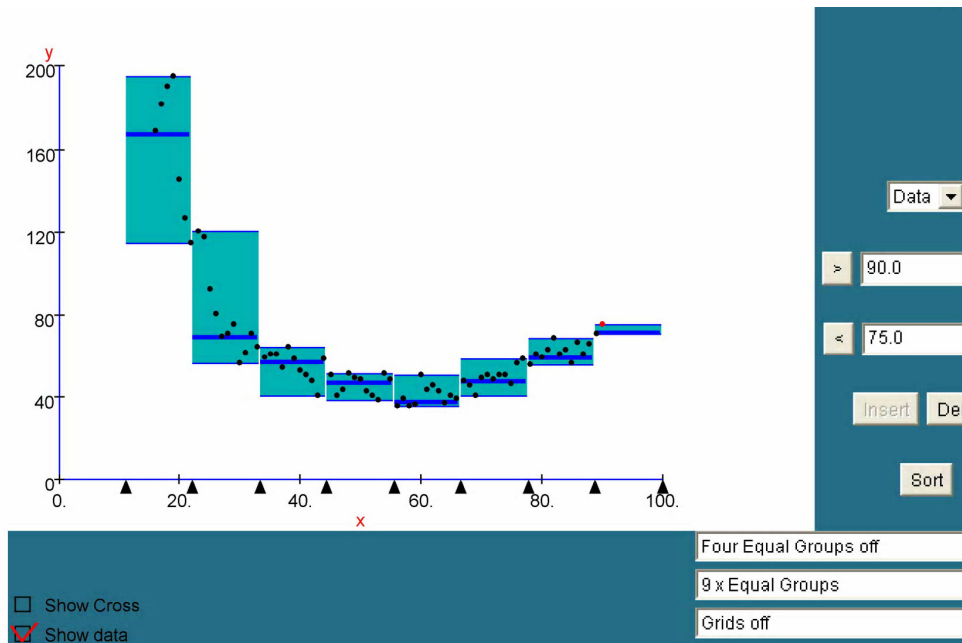
GRÁFICA 5.28.

FOUR EQUAL GROUPS OFF 4 X EQUAL GROUPS

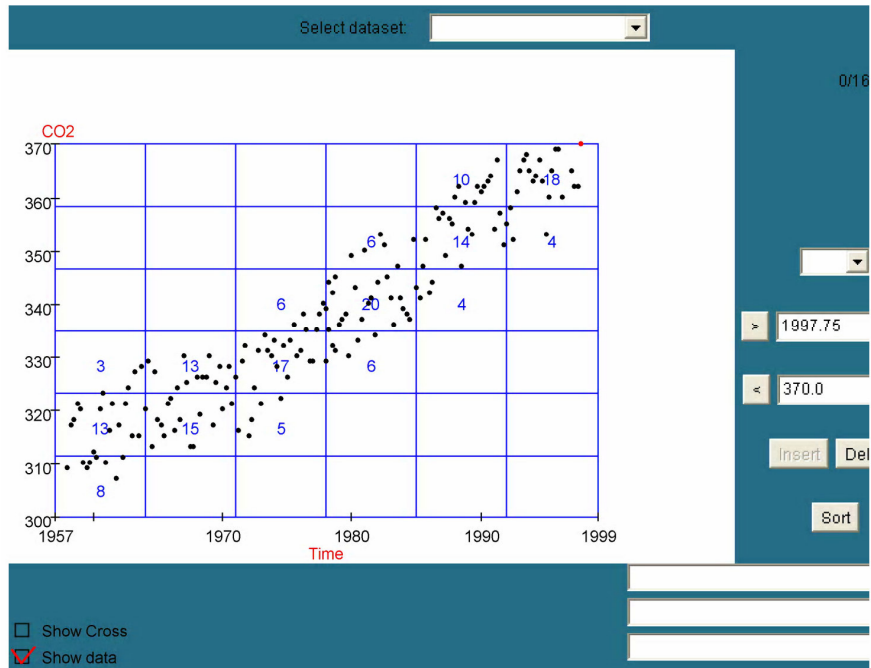


GRÁFICA 5.29..

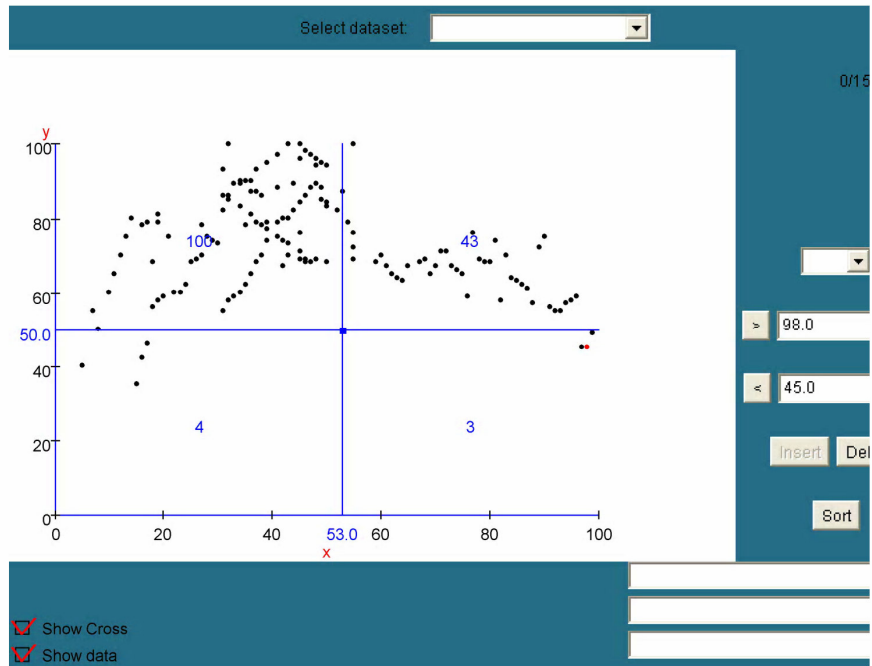
TWO EQUAL GROUPS OFF X 9 EQUAL GROUPS.



GRÁFICA 5.30.  
CUADRÍCULA DE 6X6



GRÁFICA 5.31.  
SISTEMA DE COORDENADAS



Para continuar se diseñaran situaciones problema, que tengan que ver con los conjuntos de datos dados en la herramienta, los cuales servirán como modelo para trabajar con el programa en cualquier contexto, el trabajo de los maestros será a partir de éstas situaciones crear unas nuevas y así lograr que el estudiante construya algunos de los conceptos importantes de la estadística, llegue a analizar, y a conjeturar generalidades acerca de conjuntos de datos.

En primer lugar se deja que el estudiante explore el programa por si solo, encuentre características acerca de sus funciones, y llegue a hallar algunas conjeturas acerca del uso de ésta herramienta.

### **5.8.2. Actividades**

#### *Minitool 1*

#### ***Situación problema***

Se hizo un estudio en Bogotá acerca de los casos de personas que sufren de migraña, esta enfermedad es causada por estrés, mala alimentación, esfuerzo de la vista, el medio ambiente y otras más, de esta forma el volumen de población con dicha enfermedad es bastante alto. Un laboratorio farmacéutico sacó al mercado dos clases de medicinas para esta enfermedad. Se hizo un estudio a una muestra de 40 personas que usaron dichos medicamentos, y en la siguiente gráfica se representa de color verde el tiempo de efecto del medicamento de “tipo A”, y de color rosado el del medicamento de “tipo B”. De esta forma se obtiene la gráfica 5.32.

A partir de la gráfica y utilizando las herramientas que le proporciona el programa conteste las siguientes preguntas, acerca de la eficacia de los medicamentos. Cada barra representa el tiempo de efecto de las medicinas.

1. De un título para el estudio, teniendo en cuenta todos los factores estudiados.
2. ¿Que variable se esta estudiando? ¿De que tipo es ésta?

3. ¿Cuál fue el menor y la mayor tiempo de efecto del remedio de “tipo A” aplicado?
4. ¿Cuál fue el menor y la mayor tiempo de efecto del remedio de “tipo B” aplicado?
5. Haga un estudio sobre el tiempo de efecto del medicamento de “tipo A”.  
 ¿Cuántas personas se recuperaron en menos de 40 minutos?  
 ¿Cuántas personas se recuperaron en más de 40 minutos y en menos de 80 minutos?  
 ¿Cuántas personas se recuperaron pasados 80 minutos y en menos de 120 minutos?  
 ¿Cuántas personas se recuperaron pasados 120 minutos y antes de 160 minutos?  
 ¿Cuántas personas se recuperaron pasados 160 minutos y antes de 200 minutos?
6. Haga el mismo estudio anterior, pero con la población que ingirió el medicamento de “tipo B”.
7. De acuerdo a los dos ítems anteriores ¿Cuál medicamento es menos eficaz?
8. Luego de un tiempo, el laboratorio encuentra la fórmula para otra medicina, “tipo C” ésta se prueba en una muestra de 10 personas y se muestran los resultados en la tabla 5.7.

GRÁFICA 5.32.

REPRESENTACIÓN DE SITUACIÓN PROBLEMA 1

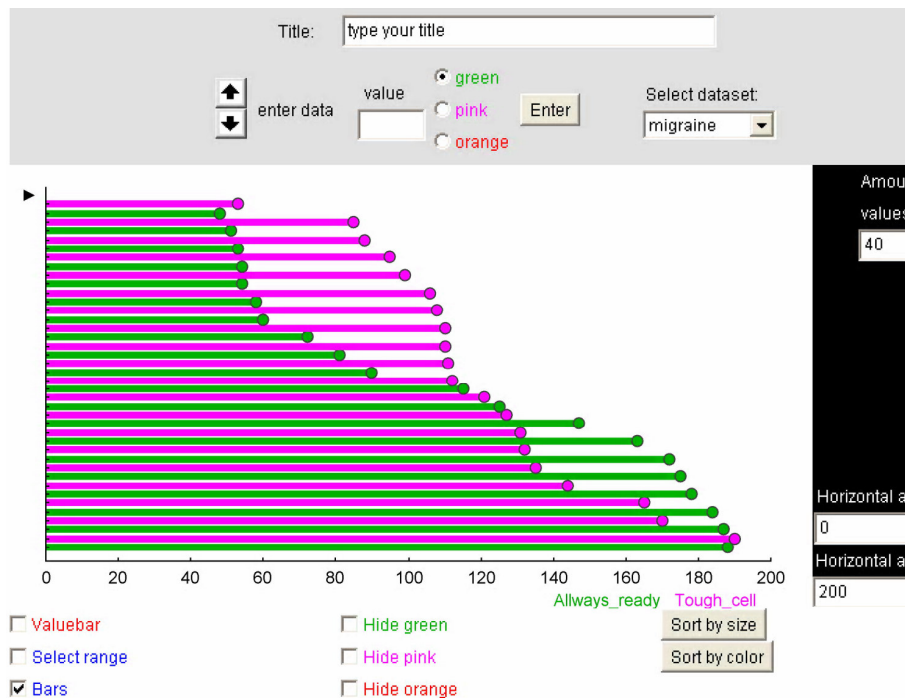


TABLA 5.8.  
EFECTO MEDICINA “TIPO C” EN 10 PERSONAS

Individuo	Tiempo de efecto (minutos)
1	35
2	16
3	5
4	17
5	100
6	87
7	36
8	91
9	58
10	157

Agregue estos datos en la gráfica de barras. Es necesario hacer variaciones en el color de estos últimos datos, ya que se podría prestar para confusiones (Gráfica 5.33).

Organice los datos y conteste las siguientes preguntas

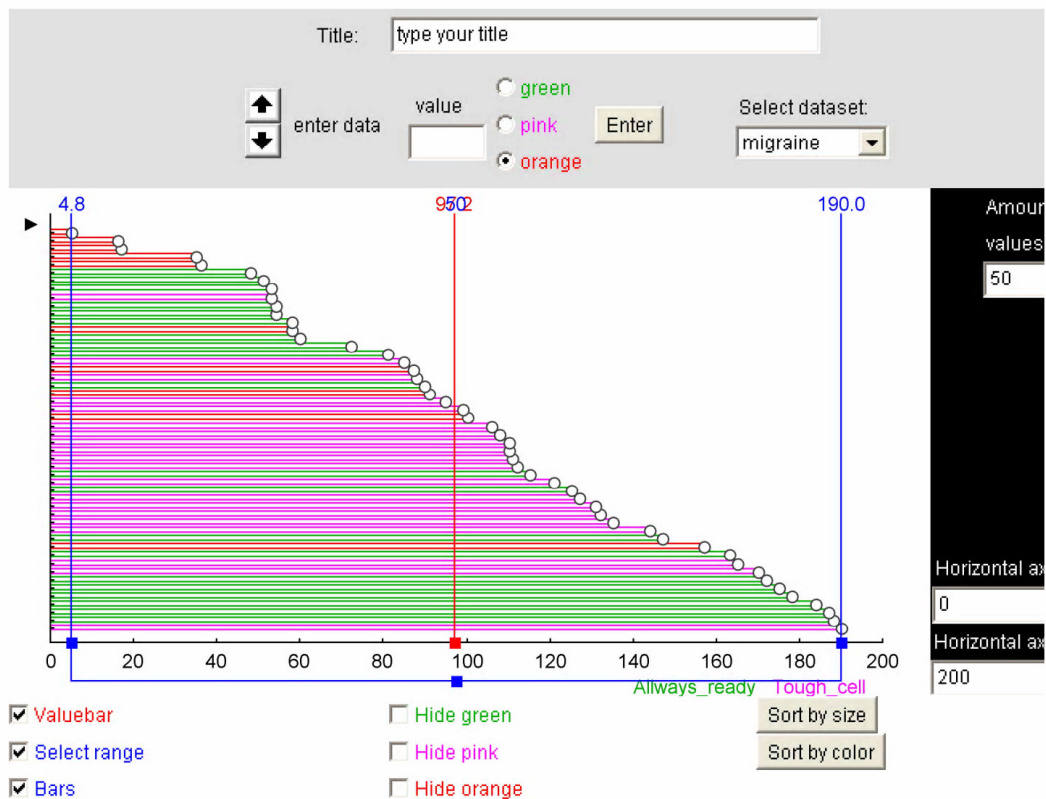
9. ¿Cuál es el tiempo de efecto de los tres medicamentos que es mas frecuente? ¿Qué nos indica este valor?
10. ¿Cuál es la moda de las frecuencias?
11. ¿Cuál es el valor medio del tiempo de efecto del medicamento de “tipo A” del “tipo B” y del “tipo C”?
12. ¿Cuál es el valor medio del tiempo de efecto de los tres tipos de medicamentos?
13. ¿Qué diferencia hay entre la mediana de las poblaciones por separado y la población total? ¿A que conclusión se puede llegar acerca de las características de los conjuntos de datos?
14. ¿En que intervalo de tiempo (tomando cada uno de no más de 20 de minutos) se encuentra una mayor cantidad de personas a las cuales les hizo efecto el medicamento de”tipo A”? ¿el medicamento de “tipo B”? Y ¿el medicamento de



“tipo C”? ¿Cuántas se encuentran en dicho intervalo? ¿El número de personas es significativo?

15. ¿En que intervalo de tiempo (no mas de 20 de minutos) se encuentran mas personas a las cuales les hizo efecto uno de los tres medicamentos?
16. Si le pidieran tomar una decisión acerca de que medicamento lanzar al mercado por su eficacia. ¿Cuál escogería? ¿Por qué?
17. Si por el contrario tuviese la oportunidad de lanzar los tres medicamentos al mercado, pero con diferentes condiciones y precios. ¿Cómo haría la selección?

GRÁFICA 5.33.  
SITUACIÓN PROBLEMA I  
ACTIVIDAD 12



## *Minitool 2*

### ***Situación problema 2***

Se hizo una encuesta acerca de la cantidad de días que las personas llevan reciclando, a partir de una noticia dada 60 días atrás sobre la importancia de este ejercicio, ya que de no hacerlo, es posible quedar sin árboles, ya que la magnitud de la tala ha venido aumentando en los últimos años en la necesidad de hacer papel y demás. Además la capa de ozono se ha venido deteriorando a causa de los gases generados por la basura que no está siendo bien distribuida.

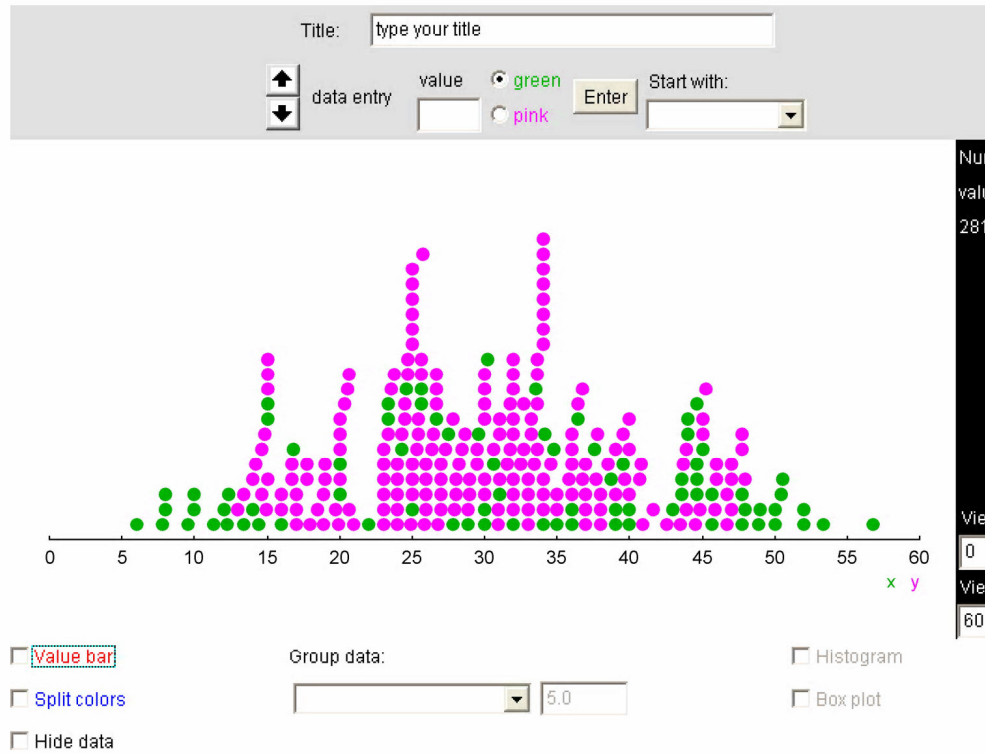
Luego de la encuesta se pudo establecer que había dos tipos de poblaciones principalmente, los que reciclan solo papel (representados con el color rosado), y los que reciclan todos los elementos (orgánicos, plástico, papel) o al menos dos de ellos (representados de color verde). (Gráfica 5.34)

### *Actividad 2*

1. Dar un título adecuado a la gráfica, que encierre todos los aspectos tenidos en cuenta para la elaboración de ésta.
2. Divida las dos poblaciones dadas, luego haga sus propios grupos de datos, para hallar en que clase se encuentra más concentrada la población.
3. Tomando un intervalo no mayor a 5 días. ¿Cuál es el intervalo en número de días mas frecuente en que las personas han reciclado papel? ¿Cuál es el intervalo de número de días en que las personas han reciclado más de un elemento? Sin tener en cuenta lo que se haya reciclado ¿Cuál es el intervalo de número de días en que las personas más han reciclado?
4. Haciendo cada vez menor el intervalo. ¿Es posible hallar el número exacto de días en que las personas han hecho más reciclaje?
5. ¿Será posible afirmar que los datos conforman una distribución normal? ¿Por qué?
6. En que población hay una mayor variabilidad respecto a los días de reciclaje en la que recicla solo papel o en la que recicla mas de un elemento. ¿Qué te podría indicar este hecho?

GRÁFICA 5.34.

SITUACIÓN PROBLEMA 2.

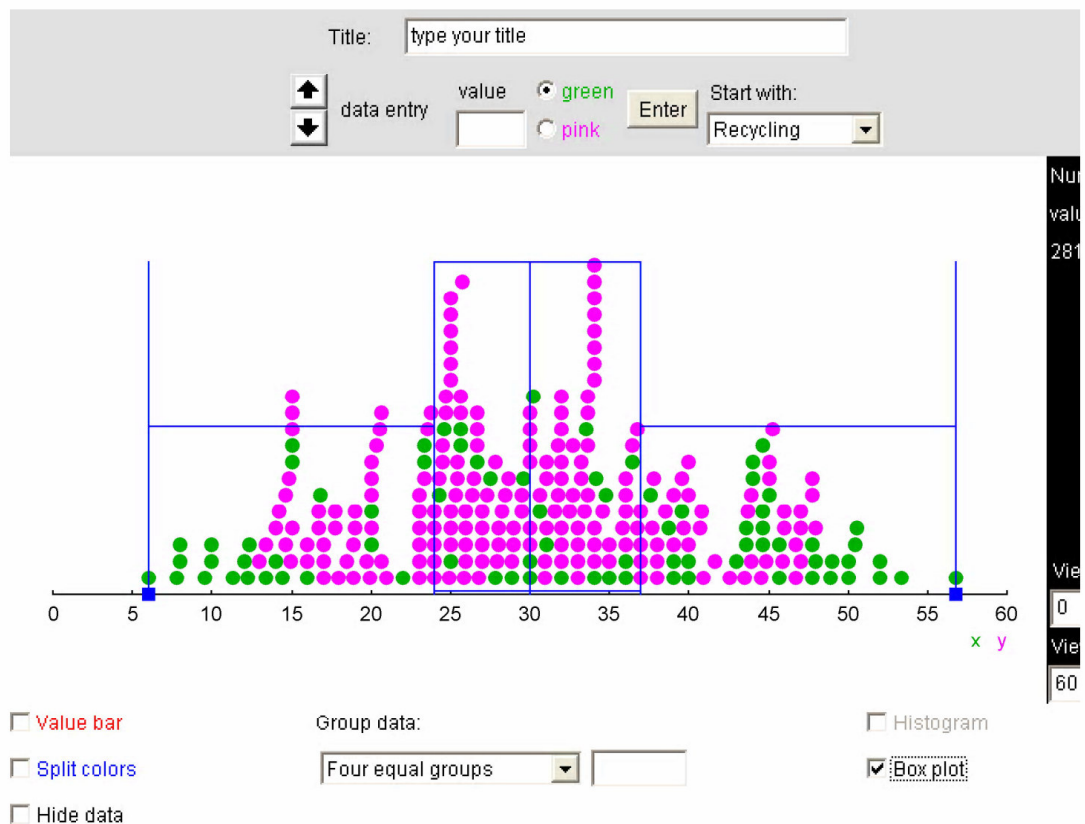


7. ¿Cuál es la mediana del conjunto de datos, de las personas que reciclan solo papel?  
¿Cuál es la mediana del conjunto de datos, de las personas que reciclan más de un elemento? ¿Cuál es la mediana de la cantidad total de las personas que reciclan?  
¿La mediana varía? ¿Por qué se presenta este hecho?
8. Divide los datos en cuatro grupos, que contengan la misma frecuencia de población.  
¿Cuál es el porcentaje de población que se encuentra en cada uno de los grupos?
9. ¿Cuál es el 25% de la población de la población? ¿Cuál es el 50% de la población de la población? ¿Cuál es el 75% de la población de la población? ¿Cuál es el 100% de la población de la población?
10. Genera una caja de puntos. ¿Cómo podrías definir este tipo de gráfica?
11. ¿Cuál es el intervalo de menor tamaño obtenido? ¿Qué indica éste?

12. Genera una caja de puntos, para las personas que reciclan solo papel y otra para las que reciclan más de un elemento. ¿Qué diferencia hay entre ambas? ¿Qué podrías concluir?
13. Es necesario escribir una noticia sobre el informe acerca del reciclaje, a partir de los datos conseguidos por medio de las gráficas y análisis redacta la información necesaria, y los consejos para llevar a cabo una mejor tarea con el reciclaje.

GRÁFICA 5.35.

SITUACIÓN PROBLEMA 2. ACTIVIDAD 11.



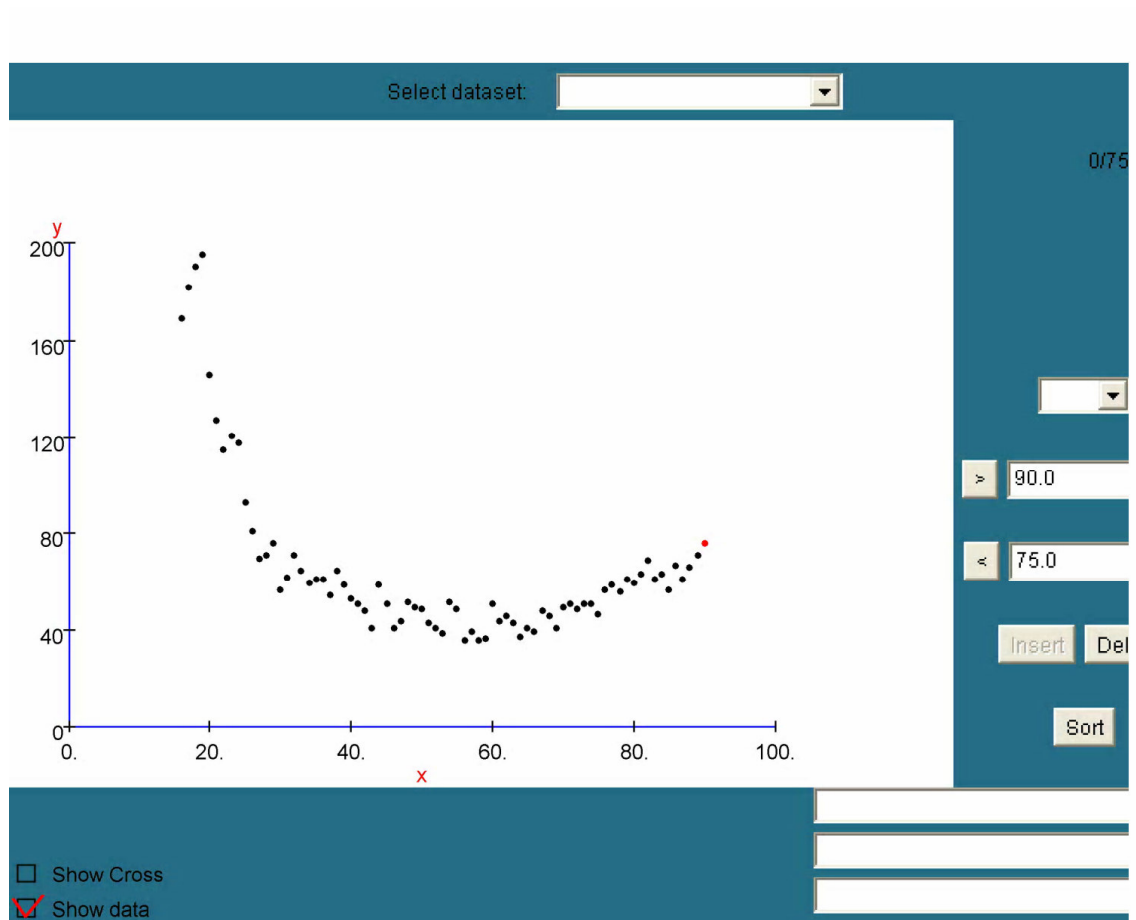
### *Minitool 3*

#### ***Situación problema 3***

En Colombia, se hizo un estudio acerca de la relación entre los accidentes que sufren las personas y la edad de estas. De éste se sacó como conjetura la siguiente gráfica. Contesta las siguientes preguntas manipulando el diagrama con ayuda del programa y las herramientas que éste facilita.

GRÁFICA 5.36.

SITUACIÓN PROBLEMA 3.

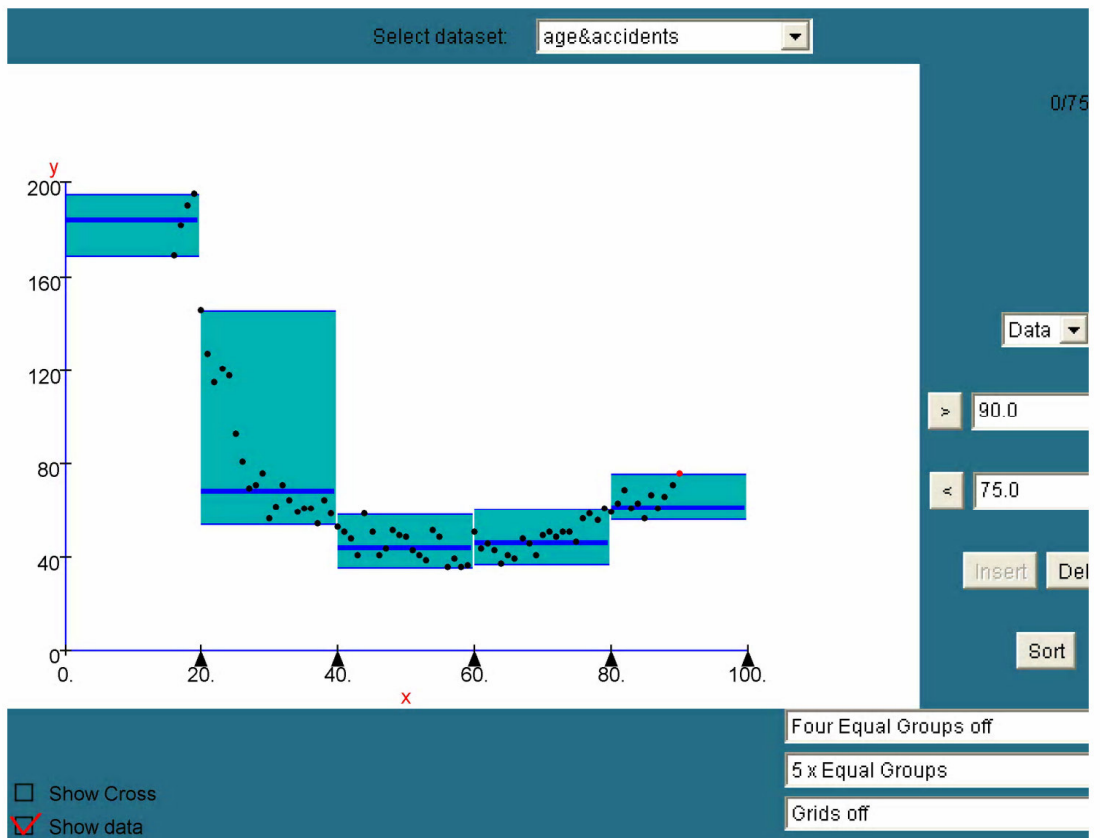


#### ***Actividad 3***

1. ¿Que variable se está representando en el eje horizontal y cuál en el eje vertical?  
Justifique tu respuesta.

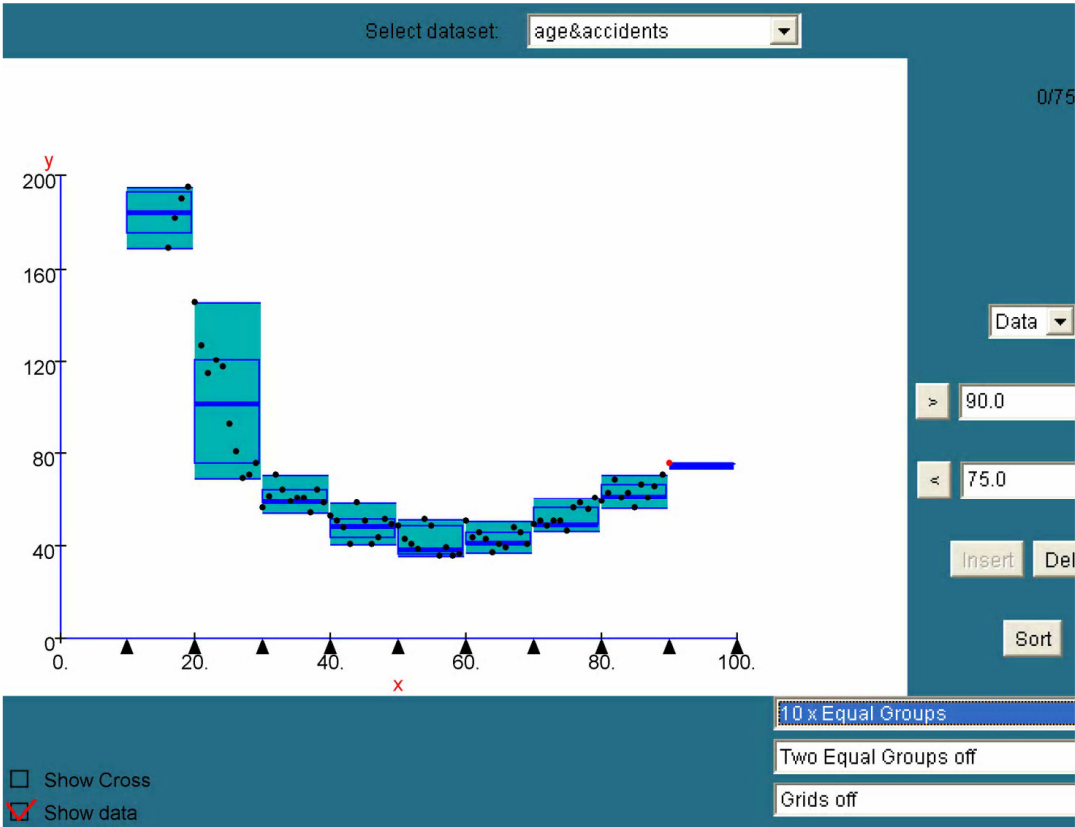
2. Diga una generalidad que pueda identificar al observar la gráfica 5.36. sin ningún cambio.
3. ¿En que intervalo de edades se presenta el mayor número de accidentes? Para responder esta pregunta divida su gráfica con una cuadrícula 5x5
4. Tomando el intervalo de 0 a 50 años y de 50 a 100 años. ¿En cuál de ellos se presenta la mayor cantidad de accidentes?
5. Tomando el intervalo de 0 a 100 accidentes y de 100 a 200. ¿Qué intervalo de accidentes es más frecuente?

GRÁFICA 5.37  
SITUACIÓN PROBLEMA 3.



Es posible que el lector esté en la capacidad de elaborar preguntas para que el estudiante extrapole acerca de los datos utilizando las gráficas 6.18 y 6.19.

GRÁFICA 5.38  
SITUACIÓN PROBLEMA 3.1



## CONCLUSIONES Y COMENTARIOS FINALES

En primer lugar, la indagación bibliográfica relacionada con el campo de investigación que atañe al tema de la noción de distribución estadística es bastante amplia y compleja. Debido a ello se debe reconocer que la elaboración de una propuesta didáctica en torno a toda la temática que se esboza en el marco conceptual fue demasiado ambiciosa, ya que a la luz del análisis de contenido debía dar cuenta de una parte descriptiva y de una parte probabilística.

A pesar de ello, considero que la propuesta presentada da buena cuenta de los aspectos de la noción de distribución estadística desde la perspectiva de la estadística descriptiva; en particular, en lo que tiene que ver con conceptos asociados a dicha noción como son: la presencia de variables, gráficas y tipos de tabulaciones de frecuencia considerados para el estudio de la construcción y/o representación de una distribución de frecuencias.

En cuanto al análisis de textos, vale la pena señalar que el formato de los libros de texto hoy en día, permite una formación más amplia del estudiante en el campo de la matemática, porque no sólo se centra en el concepto, sino que al tratar la historia y las diversas aplicaciones de los temas en estudio, se le da también importancia al origen y desarrollo de los conceptos estadísticos, como al uso que se le puede dar a éstos dentro de los diversos contextos.

Tomando como referencia el análisis del aprendizaje, se concluye que el razonamiento estadístico se puede definir como la manera en la que los sujetos utilizan las ideas estadísticas, es decir dan un sentido a la información obtenida luego de llevar a cabo un proceso determinado de análisis. Por otra parte el estudio de los errores que cometen los estudiantes nos proporcionan resultados útiles, ya que las ideas presentadas en el trabajo



respecto a este tema, permiten inferir que es posible hallar regularidades y con estas llegar a construir modelos, por ello, éste resulta ser un factor que todo docente debería considerar.

La elaboración de esta monografía permite evidenciar no sólo la importancia de presentar situaciones reales o que parezcan reales en la enseñanza de la estadística sino también la existencia de una gran variedad de contextos en los que se encuentra presente dicha área, los cuales podrían ser utilizados para despertar el interés de los estudiantes. La propuesta hace evidente que es posible generar actividades en el aula de clase basadas en estudios que se llevan a cabo casi diariamente alrededor del educando.

Finalmente, la exploración de software educativo en torno a esta temática, permite llevar a cabo diseños de actividades en las que el estudiante llegue a generar alguna clase de conocimiento, en la que el docente también debe asumir su reto de investigador.

## REFERENCIAS BIBLIOGRAFICAS

Ausubel, D. P., Novak, J. D. y Hanesian, H. (1983). *Psicología educativa. Un punto de vista cognoscitivo*. (México: Trillas).

Baena, J. (1999). Sistemas de datos en el currículo. *Uno*, (20), 9 – 24

Barr, G. V. (1980). Some student ideas on the median and the mode. *Teaching Statistics*, 2 (2), 38-41.

Batanero, C., Estepa A. & Godino J. (1991). Análisis exploratorio de datos: Sus posibilidades en la enseñanza secundaria. *Suma*, (6).

Batanero C., Rodino J., Green D., Holmes P., & Vallecillos A. (1996) Errores y dificultades en la comprensión de los conceptos estadísticos elementales. *Internation Journal of Mathematics Education in Science and Technology*, 25(4), 527-547.

Beltrán, L., Rodríguez, B. & Dimaté, M. *Matemáticas Prentice Hall 6* (pp. 236 – 271). Bogotá, Colombia: Pearson Educación de Colombia Ltda.

Beltrán, L., Rodríguez, B. & Dimaté, M. *Matemáticas Prentice Hall 7* (pp. 317 – 349). Bogotá, Colombia: Pearson Educación de Colombia Ltda.

Beltrán, L., Rodríguez, B. & Dimaté, M. *Matemáticas Prentice Hall 9* (pp. 306 – 355). Bogotá, Colombia: Pearson Educación de Colombia Ltda.

Beltrán, L., Rodríguez, B. & Dimaté, M. *Matemáticas Prentice Hall 10* (pp. 310 – 355). Bogotá, Colombia: Pearson Educación de Colombia Ltda.

Camargo, L., Garcia, G., Leguizamón, C., Samper, C. & Serrano, C. (2003). *Alfa 6 con estándares: Serie de Matemáticas para educación secundaria y media* (pp. 281 – 309). Bogotá, Colombia: ed. Norma.

Camargo, L., Garcia, G., Leguizamón, C., Samper, C. & Serrano, C. (2003). *Alfa 8 con estándares: Serie de Matemáticas para educación secundaria y media* (pp. 312– 349). Bogotá, Colombia: ed. Norma.

Canavos, G. (1988). *Probabilidad y Estadística*. México: McGraw Hill.

Castañeda, J. P., y Rodrigo, M. J. (1993). Modelos mentales en el razonamiento predictivo de expertos. *Revista de Psicología General y Aplicada*, 46 (3), 275-283.

Chou, Ya-Lun. (1972). *Análisis Estadístico*. México: Nueva editorial Interamericana.

Confrey, J. (1990). A review of the research in student's conceptions in mathematics, science and programming. *Review of Research in Education*. 16, 3-35.

Croxton F. & Cowden D. (1992). *Estadística general aplicada*. México: Fondo de Cultura económica.

Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, 18 (5), 382-393.

Domínguez, J. (1999). Información y análisis exploratorio de datos. *Uno*, (20), 41 – 56.

Estepa Castro, A. (1990). *Enseñanza de la estadística basada en el uso de ordenadores: Un Estudio exploratorio*. Memoria de Tercer Ciclo. (Universidad de Granada: Departamento de Didáctica de la Matemática).

Estévez, C. (1996). *Evaluación integral por procesos. Una experiencia construida desde y en el aula*. Cooperativa Editorial Magisterio. Colombia.

Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. (Dordrecht: D. Reidel).

Fong, G.T., Krantz, D.H., y Nisbett, R. E. (1986). The Effects of Statistical Training on Thinking about Everyday Problems. *Cognitive Psychology*, 18, 253-292.

García, F. & Garzo, F. (1988). *Estadística*. España: McGraw Hill.

García, V. (1995). *Enseñanza de las matemáticas en la educación secundaria*. Madrid: Ediciones RIALP, S.A.

Garfield, J. (1998). The statistical Reasoning Assessment: development and validation of a research tool. *Proceedings of the Fifth International Conference on Teaching of Statistics*, 781-786.

Garfield, J. y Alhgren, A.(1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education*, 19 (1), 44-63.

Garfield, J., y delMas, R. (1989). Reasoning about chance events: Assessing and changing students' conception of probability. In C. Maher, G. Goldin, y B. Davis (eds.), *The Proceedings of the Eleven Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (vol. II, pp. 189-195). Rutgers: Rutgers University Press.

Gilbert, N. (1980). *Estadística*. México: Nueva editorial interamericana.

Godino, J., Batanero C. & Cañizares M. (1996). *Azar y probabilidad, fundamentos didácticos y propuestas curriculares*. España: Editorial Síntesis.

Goodchild, S. (1988). School pupil's understanding of average. *Teaching Statistics*, 10 (3), 77-81.

Green, D. R. (1992). Data analysis: what research do we need?. Paper presented at the *Eighth ISI Round Table Conference on Teaching Statistics*. Lennoxville, 1992.

Guerrero, S. (1999). Tratamiento de la información. *Uno*, (20), 5 – 8.

Haber, A. (1973). *Estadística General*. Los Ángeles California: Fondo Educativo Interamericano, S.A.

Hawkins, A. y Kapadia, R. (1984). Children's conceptions of probability - a psychological and pedagogical review. *Educational Studies in Mathematics* 15, 349-377.

Holmes, P. y cols.(1980). *Statistics in your world*. (Slough: Foulsham Educational).

Huck, S.; Cross, T.L.; y Clark, 5. B. (1986). Overcoming misconceptions about z-scores. *Teaching Statistics*, 8 (2), 3 8-40.

Kahneman, D.; Slovic, P. y Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. (New York: Cambridge University Press).

Konold, C. (1995). Issues in Assessing Conceptual Understanding in Probability and Statistics. *Journal of Statistics Education*, 3(1). (on-line).

Landwehr, J. y Watkins, A. E.(1 986). *Exploring data*. (Palo Alto: Dale Seymour).

Li, K. Y. y Shen, S. M. (1992). Student's weaknesses in statistical projects. *Teaching Statistics*, 14 (1), 2-8.

Loosen, F.; Lioen, M. y Lacante, M. (1985). The standard deviation: some drawbacks of an intuitive approach. *Teaching Statistics*, 7 (1), 2-5.

Ministerio de Educación Nacional (1985). *La estadística: Conceptos, importancia y aplicaciones*. Colombia: Oficina sectorial de planeación educativa: División de estadística y sistemas.

Ministerio de Educación Nacional (1998). *Lineamientos Curriculares. Áreas obligatorias y fundamentales*. Colombia: Magisterio.

Moreno, V. & Restrepo, M. (2003). *Alfa 7 con estándares: Con énfasis en competencias* (pp. 240 – 274). Bogotá, Colombia: ED. Norma.

Moreno, V. & Restrepo, M. (2003). *Alfa 10 con estándares: Con énfasis en competencias* (pp. 278 – 311). Bogotá, Colombia: ED. Norma.

Moreno, V. & Restrepo, M. (2003). *Alfa 11 con estándares: Con énfasis en competencias* (pp. 266 – 300). Bogotá, Colombia: ED. Norma.

Moya, Marcos. (1995). *Probabilidad y Estadística un enfoque teórico práctico*. Costa Rica: Editorial tecnológica de Costa Rica

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

Nisbett, R.E., Krantz, D.H., Jepson, C., y Kunda, Z. (1983). The Use of Statistical Heuristics in Everyday Inductive Reasoning. *Psychological Review*, 90(4), 339-363.

- Nisbett, R.E., y Ross, L. (1980). *Human Inference: Strategies and Shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall
- Peña, D. (1986). *Estadística: Modelos y métodos*. Madrid: Alianza editorial S.A.
- Pollatsek, A. Lima, S. y Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics*, 12, 191-204.
- Quijano, M. (1988). *Iniciación a la estadística, la probabilidad y la combinatoria para la enseñanza en educación básica*. Bogotá: Pontificia Universidad Javeriana facultad de ciencias sociales y educación.
- Radatz, H. C. (1980). Student errors in the mathematical learning: a survey. *For the Learning of Mathematics*, 1(1), 16-20.
- Rivadulla, A. (1991). *Probabilidad e inferencia científica* (Barcelona: Anthropos).
- Robert, Jonson. (1990). *Estadística Elemental*. México: Ibero América.
- Russell, S. J. y Mokros, J. R. (1991). What's typical?: Children's ideas about average. In: D. Vere-Jones (Ed.) *Proceedings of the Third International Conference on Teaching Statistics* (Voorburg, The Netherlands: International Statistical Institute), 307-313.
- Scholz, R. (1987). *Decision making under uncertainty*. Amsterdam. North Holland.
- Scholz, R. (1991). Psychological research on the probability concept and its acquisition. In R. Kapadia (Ed.) *Chance Encounters: Probability in Education*. (Amsterdam: Reidel), 213-249.

Schuyten, G. (1991). Statistical thinking in psychology and education. In: D. Vere-Jones (Ed.) *Proceeding of the Third International Conference on Teaching Statistics* (Voorburg, The Netherlands: International Statistical Institute), 486-490.

Shaughnessy, J. M. (1983). The Psychology of Inference and the Teaching of Probability and Statistics: Two Sides of the Same Coin? En R. W. Scholz (ed.), *Decision Making Under Uncertainty*, 325-350. North-Holland: Elsevier Science Publishers B.V.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In A. Grouws (Ed.). *Handbook of research in teaching and learning mathematics* (New York: McMillan), 465-494.

Spiegel, Murray. (1998). *Estadística*. Colombia. Mc Graw Hill.

Steinbring, H. (1990). The nature of stochastic knowledge and the traditional mathematics curriculum- some experience with in-service training and developing material. En: A. Hawkins (Ed.): *Training Teachers to Teach Statistics*. (Voorburg, The Netherlands:I.S.I.), 2-19.

Strauss, S. y Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education*, 19 (1), 64-80.

Well, A., Pollatsek, A., y Boyce, S. (1990). Understanding the effects of sample size on the variability of the mean. *Organizational behavior and human decision processes*, 47, 289-312.

Wert, J. (1938). *Educational Statistics*. New York: McGraw Hill

Yamane, Taro. (1973). *Estadística*. México: Harla.