ASIGNATURA TRANSVERSAL DE LA FACULTAD DE CIENCIAS BÁSICAS E INGENIERÍA

ASIGNATURA: Estadística Descriptiva

CORPORACIÓN UNIVERSITARIA REMINGTON

DIRECCIÓN PEDAGÓGICA

2011

CRÉDITOS

El módulo de estudio de la asignatura Estadística Descriptiva Asignatura Transversal de la Escuela de Ciencias Básicas e Ingeniería es propiedad de la Corporación Universitaria Remington. Las imágenes fueron tomadas de diferentes fuentes que se relacionan en los derechos de autor y las citas en la bibliografía. El contenido del módulo está protegido por las leyes de derechos de autor que rigen al país.

Este material tiene fines educativos y no puede usarse con propósitos económicos o comerciales.

AUTOR

Ángela Quintero Echeverri
Tecnóloga en Sistematización de datos. Estudiante de último semestre de Psicología, Coordinadora académica de la Escuela de Ciencias Básicas e Ingeniería, Profesora de la Corporación Universitaria Remington
rosa.quintero@remington.edu.co

Nota: el autor certificó (de manera verbal o escrita) No haber incurrido en fraude científico, plagio o vicios de autoría; en caso contrario eximió de toda responsabilidad a la Corporación Universitaria Remington, y se declaró como el único responsable.

RESPONSABLES

ESCUELA DE CIENCIAS BÁSICAS E INGENIERÍA
Director Dr. Mauricio Sepúlveda

Director Pedagógico
Octavio Toro Chica
dirpedagogica.director@remington.edu.co

Coordinadora de Medios y Mediaciones
Angélica Ricaurte Avendaño
mediaciones.coordinador01@remington.edu.co

GRUPO DE APOYO

Personal de la Unidad de Medios y Mediaciones
EDICIÓN Y MONTAJE
Primera versión. Febrero de 2011.

TABLA DE CONTENIDO

1. INTRODUCCIÓN

La estadística es la ciencia de los datos, por tanto, cuando se aplica el método estadístico, se recolectan, se sintetizan, se organizan, se analizan y se interpretan los datos.

La estadística descriptiva se encarga de describir los datos por medio de tablas, gráficos y medidas; en este módulo se explicará cómo lograrlo. Se pretende que el estudiante aplicando paulatinamente cada paso que se explica, lo logre.

Para que se llegue al objetivo terminal: Estudiar métodos de organización, análisis y presentación de un conjunto de datos asociados a una situación problémica por medio del modelo de representación estadístico caracterizando un conjunto de datos a partir de mediciones estadísticas para la obtención de conclusiones que sirvan de apoyo en la toma de decisiones, se ha diseñado el módulo de una forma especial e innovadora; en la primera parte se definen los conceptos generales que se requieren en estadística y posteriormente se han dividido las unidades de acuerdo con los tipos de datos, a saber:

- ✓ Datos cualitativos.
- ✓ Datos cuantitativos ordenados en fila.
- ✓ Datos cuantitativos agrupados en intervalos.

Nota: En cada tipo de datos, se explica cómo se pueden identificar, recolectar, organizar y describir por medio de tablas gráficos y medidas para finalmente llegar a conclusiones y basado en éstas, tomar las decisiones correspondientes.

El módulo está construido con un lenguaje sencillo, con ejercicios aplicados a la cotidianidad y a situaciones prácticas inherentes al plan de estudio del estudiante de La Corporación Universitaria Remington de cualquier programa existente en la universidad, con el fin de que, de una forma pedagógica, se aprenda y se logren los objetivos.

La Estadística Descriptiva le proporciona al profesional las herramientas que necesita para recoger, organizar, analizar y presentar datos con el fin de tomar decisiones.

Debido a que el estudiante, de educación a distancia de La Corporación Universitaria Remington, requiere un método de aprendizaje de forma tal que el profesor sea un orientador y que dicho estudiante sea autogenerador de su conocimiento, obviamente con la asesoría del docente, se ha creado este módulo.
Este módulo está diseñado con un lenguaje sencillo y con ejercicios que son aplicados a la cotidianidad del estudiante, a su entorno social y laboral, pues de esta forma podrá realizar

investigaciones estadísticas en un futuro, ya sea a corto plazo, en otras asignaturas, o a largo plazo cuando esté realizando su labor como profesional.

2. MAPA DE LA ASIGNATURA

ESTADÍSTICA DESCRIPTIVA

PROPÓSITO GENERAL DEL MÓDULO

Recolectar, organizar, resumir y analizar datos con el fin de tomar decisiones frente a la incertidumbre. Se debe hacer la descripción de los datos por medio de gráficas, tablas y medidas.

OBJETIVO GENERAL

Analizar un conjunto de datos, partiendo de mediciones estadísticas, obteniendo conclusiones que sirvan de apoyo en la toma de decisiones.

OBJETIVOS ESPECÍFICOS

- Explorar los conceptos generales de la estadística, analizando datos cualitativos.
- Describir por medio de tablas, gráficos y medidas los datos ordenados en fila.
- Analizar datos cuantitativos agrupados en intervalos de clase.

UNIDAD 1

Fundamentación conceptual de términos. Representa gráficamente datos cualitativos. Interpreta medidas

UNIDAD 2

Representa gráficamente datos ordenados en fila. Interpreta medidas de datos ordenados en fila.

UNIDAD 3

Representa gráficamente datos agrupados. Interpreta medidas de datos agrupados

3. UNIDADES

3.1. Conceptos generales y datos cualitativos

Objetivo General

Explorar los conceptos generales de la estadística, analizando datos cualitativos.

Objetivos específicos

- Explicar los conceptos generales de Estadística Descriptiva.
- Realizar la ficha técnica.
- Construir la tabla de distribución de frecuencias e interpretarla.
- Realizar los gráficos para datos cualitativos e interpretarlos.

Prueba Inicial

Responde falso o verdadero a las siguientes preguntas:

- La estadística descriptiva describe los datos por medio de gráficos, tablas y medidas.
- Una población estadística se refiere a las personas que habitan en un lugar.
- La muestra es cualquier parte de la población.
- Cuando se clasifican datos sobre el sexo de los estudiantes, se refiere a una variable cualitativa.
- El número de hijos es una variable cuantitativa continua.
- La estatura es una variable cuantitativa continua.

3.2. CONCEPTOS GENERALES

3.2.1. Estadística

Es una ciencia auxiliar y la compilación, organización, resumen, presentación y análisis de datos numéricos cuya función principal es elaborar principios y métodos que nos ayuden a tomar decisiones frente a la incertidumbre.

Se puede aplicar en todas las áreas de investigación bien sea técnica, tecnológica, científica. Ejemplo: análisis comparativos de ingresos y egresos, análisis de mercado para la introducción de un nuevo producto.

1. **ESTADÍSTICA DESCRIPTIVA**

Es la presentación en forma comprensible de los datos recolectados, puede ser en forma tabular, gráfica o numérica.

2. **ESTADÍSTICA INFERENCIAL**

"Se puede definir como aquellos métodos que hacen posible la estimación de la población o la toma de una decisión referente a una población, basándose sólo en los resultados de una muestra" (Berenson y levine, 1996, p.3). Es la rama de la estadística que trata de los procesos y comprensión de la teoría de estimación y prueba de hipótesis.

3. **POBLACIÓN**

Es la totalidad de unidades elementales sobre los cuales se desea información. Ejemplos:

a) Si el contador de la empresa desea investigar las deudas que ha adquirido la compañía en el último año, la población estaría conformada por todas las cuentas por pagar del último año.
b) Si un ingeniero agrónomo desea investigar en una finca cafetera si hay Broca, la población sería todas las matas de café que hay en dicha finca.
c) Si un médico veterinario desea investigar si hay gripa aviar en una granja, la población sería todas las aves que hay en la granja.

4. **POBLACIÓN FINITA** Es una población que no es indefinidamente grande o que sólo contiene un número finito de datos.

Ejemplo: Estudiantes de la C.U.R

5. **POBLACIÓN INFINITA**

Contiene un número infinitamente grande de unidades elementales (datos de la población)
Ejemplo: Al lanzar una moneda indefinidamente, el número de caras que se pueden obtener.

6. **CENSO**

Estudio de una población. Llamado también enumeración completa.

7. **PARÁMETRO**

”Es una medida descriptiva numérica de una población” (Mendenhall y Sincich, 1997, p.39).

De los ejemplos de la población:

a) Promedio mensual de las cuentas por pagar.
b) El 5% de las matas de café tienen Broca.
c) En la granja hay 850 pollos.

8. **MUESTRA**

Es el elemento básico o parte representativa de una población. Ejemplo: si deseamos investigar el nivel académico de los estudiantes de educación a distancia de la CUR, la población está conformada por todos los estudiantes de educación a distancia de la CUR y para extraer la muestra tendríamos que seleccionar proporciones iguales de estudiantes de todos y cada uno de los sitios donde funciona la educación a distancia de la CUR; si seleccionamos estudiantes de un solo lugar, por ejemplo de Puerto Berrio, esta parte no sería representativa.

9. **MUESTREO**

Es el estudio de la muestra y la relación entre la población y la muestra tomada de ella.

10. **MUESTRA AL AZAR**

Es aquella que se extrae con la condición de que posea características de los otros elementos de la población, cada elemento de la población tiene la misma probabilidad de ser elegido. Se denomina también **muestra aleatoria;** recibe este nombre porque para seleccionarla se hace por sorteo o por números aleatorios.

11. **ESTADÍSTICA O ESTADÍGRAFO:**

"Es una medida descriptiva numérica calculada a partir de los datos de la muestra" (Mendenhall y Sincich, 1997, p.39), por ejemplo, si en vez de las poblaciones que tenemos de ejemplo, extraemos muestras, los parámetros se convierten en estadísticas.

12. **VARIABLE**

Es cualquier característica que se analiza de una población, son los datos que se estudian; pueden ser cualitativas y cuantitativas.

13. **VARIABLE CUALITATIVA**

Es la que se refiere a atributos o cualidades, se divide en categorías y no es numérica. Ejemplos: estado civil (soltero, casado, viudo, unión libre) evaluar un producto (bueno, regular, malo) equipo de fútbol colombiano que más le gusta (Millonarios, Nacional, Medellín, Once Caldas, entre otros).

14. **VARIABLE CUANTITATIVA**

Se refiere a datos numéricos. Ejemplo: ingresos mensuales, edad, número de nacimientos. Este tipo de variable se divide en dos clases:

a. VARIABLE CUANTITATIVA DISCRETA

Sólo puede tomar valores enteros. Ejemplo: número de hijos por familia, número de empleados, número de estudiantes, como podemos apreciar, los ejemplos anteriores se refieren a personas, en estos casos serán variables cuantitativas discretas, también podía ser número de goles marcados en un partido, número de carros que pasan por un peaje, número de errores cometidos en una evaluación. Estas variables se hayan por conteo.

b. VARIABLE CUANTITATIVA CONTINUA

Puede asumir cualquier valor numérico, es decir, cualquier número real. Se encuentra con medición, peso, longitud, tiempo, volumen, velocidad y temperatura. Ejemplo: tiempo que se gasta en la elaboración de una evaluación, las dimensiones de un producto, el peso de un niño al nacer; además todo lo que se refiera a dinero como: gastos mensuales, costo de un producto, sueldo de los empleados.

Ejercicios tema 1. Defina cada uno de los conceptos generales.

Enumere 5 ejemplos para cada una de las variables definidas.

3.3. Redondeo Y Ficha Técnica

3.3.1. Redondeo

Es lo mismo que aproximar, para efectos de exactitud, en Estadística, se aconseja redondear con 3 cifras decimales. El redondeo se hace así:

1. Si la cifra que sigue a la cifra de redondeo es menor que 5, la cifra de redondeo se deja igual. Ejemplo: un cálculo dio 0,333333 entonces se redondea 0,333

2. Cuando la cifra que sigue a la cifra de redondeo es igual o mayor que 5, la cifra de redondeo se aumenta en 1 Ejemplo: un cálculo dio 0,666666 entonces se redondea 0,667

FICHA TÉCNICA

Cuando se hace la estadística descriptiva se debe dar una información general sobre la investigación que se hizo. Esta información tiene:

1. **Población:** se dicen cuáles fueron las unidades elementales que se investigaron.
2. **Muestra:** Se define la muestra y el tamaño en caso de que exista. Cuando no hay muestra se dice que se hizo un censo.
3. **Descripción de la variable:** Se dice cuál es la característica de la población que se va a investigar.
4. **Tipo de variable:** Se dice si la variable es cualitativa o cuantitativa discreta o continua.

Ejemplo 1:

Entre las familias de Medellín se eligieron aleatoriamente 5.000 con el fin de investigar el número de hijos por familia.

POBLACION: familias de Medellín.
MUESTRA: 5.000 familias de Medellín elegidas al azar.
DESCRIPCION DE LA VARIABLE: número de hijos por familia.
TIPO DE VARIABLE: cuantitativa discreta.

Ejemplo 2:
Los estudiantes de sistemas de Ciudad Bolívar diseñaron un software administrativo e hicieron la demostración de él ante los empresarios de la misma ciudad, posteriormente, les dijeron que lo evaluaran entre: bueno, regular y deficiente.

POBLACION: Los empresarios de Ciudad Bolívar
MUESTRA: no hay, se hizo un censo
DESCRIPCION DE LA VARIABLE: evaluar el software administrativo
TIPO DE VARIABLE: cualitativa.

Ejercicios tema 2. Redondeo de cifras

a. **Redondeé los siguientes números**

1. 0.53689
2. 0.2674343
3. 0.0989898
4. 0.999999
5. 89.999555

b. **Elabore la ficha técnica de las siguientes investigaciones**:

1. Entre la producción del almacén "Muebles S.A." del último mes, se eligieron al azar 500 muebles para catalogarlos por su calidad entre muebles de 1ª, 2ª y defectuosos.

2. Debido a las dificultades presentadas con la nómina de la cía. "Análisis", el jefe de personal escogió al azar 100 liquidaciones de nómina entre las de los últimos 3 períodos, para analizar el número de errores cometidos por cada liquidación.

3. Un ingeniero de sistemas diseñó un software contable y luego de hacer una demostración de él ante los estudiantes del último semestre de Contaduría Pública de Corporación Universitaria Remington, los interrogó sobre cuánto estarían dispuestos a pagar por dicho software.

3.4. Datos Cualitativos

Cuando existen datos cualitativos, después de realizar la ficha técnica, se debe hacer una tabla de distribución de frecuencias que tendrá:

Xi: Que es la variable de investigación o dato que se investiga.

FRECUENCIA ABSOLUTA (*fai* o *ni*)

Es el número de veces que se repite cada clase o categoría. Al sumar todas las frecuencias absolutas, se encuentra N o n
N = indica el tamaño de la población, se utiliza cuando se hace un censo.
n = indica el tamaño de la muestra, se utiliza cuando se hace un muestreo.

FRECUENCIA RELATIVA *(fri* o *hi*)

Es la relación que existe entre la frecuencia absoluta de cada categoría y el tamaño de la muestra o población según el caso.

$$fri = \frac{fai}{n \text{ ó } N} \quad \text{Ó} \quad hi = \frac{ni}{n \text{ ó } N}$$

Debido a que esta relación es de una parte al todo, al sumar todas las frecuencias desde la primera hasta la última categoría, el resultado es igual a uno.

PORCENTAJE (%)

Si se multiplican las frecuencias relativas por cien, se encuentra el porcentaje de cada categoría y la suma de todos los porcentajes es igual a cien.

Ejemplo

Una empresa productora de software eligió aleatoriamente a un grupo de contadores de Medellín para que evaluaran un paquete de nómina. Los resultados que se encontraron fueron los siguientes:

B R R R R R B B B	Donde:	B=bueno
B R M M M R B B B		R=regular
B M R R R R B R R		M=malo
R R B M M M R R R		
R R B M B R R R R		

Esto indica que el primer contador dijo que el paquete era bueno, el segundo, regular y así sucesivamente, el último dijo que era regular. Como tenemos 45 resultados este fue el total de contadores encuestados, además como nos dicen que se escogieron aleatoriamente, es una muestra.

FICHA TECNICA
POBLACION: Grupo de contadores de Medellín
MUESTRA: 45 contadores escogidos aleatoriamente
DESCRIPCION VARIABLE: Evaluar el paquete de nómina
TIPO VARIABLE: cualitativa

TABLA DE DISTRIBUCIÓN DE FRECUENCIAS

Xi (evaluación)	**fai(contadores)**	**fri**	**%**
Bueno	13	0,289	28,9
Regular	24	0,533	53,3
Malo	8	0,178	17,8
Total	**n = 45**	**1**	**100**

Como es un muestreo, se utiliza al sumar las frecuencias absolutas **n;** si fuese un censo se hubiera puesto **N**

La primera frecuencia relativa se calculó así: 13/45 = 0,2888… por tanto se redondea a 0,289 y esta cantidad se multiplica por cien para obtener el porcentaje.

La segunda frecuencia relativa se calculó así: 24/45 = 0,53333… por tanto se redondea a 0,533 y esta cantidad se multiplica por cien para obtener el porcentaje.

La tercera frecuencia relativa se calculó así: 8/45 = 0,178888… por tanto se redondea a 0,179 y esta cantidad se multiplica por cien para obtener el porcentaje.

Interpretación de la frecuencia absoluta:

El grupo de contadores evaluaron el paquete de nómina de la siguiente forma:

13 contadores dijeron que el paquete es bueno.
24 contadores dijeron que el paquete es regular.
8 contadores dijeron que el paquete es malo.

Interpretación de la frecuencia relativa:

El grupo de contadores evaluaron el paquete de nómina de la siguiente forma:
El 28,9% de los contadores dijeron que el paquete es bueno.
El 53,3% de los contadores dijeron que el paquete es regular.
El 17,8% de los contadores dijeron que el paquete es malo.

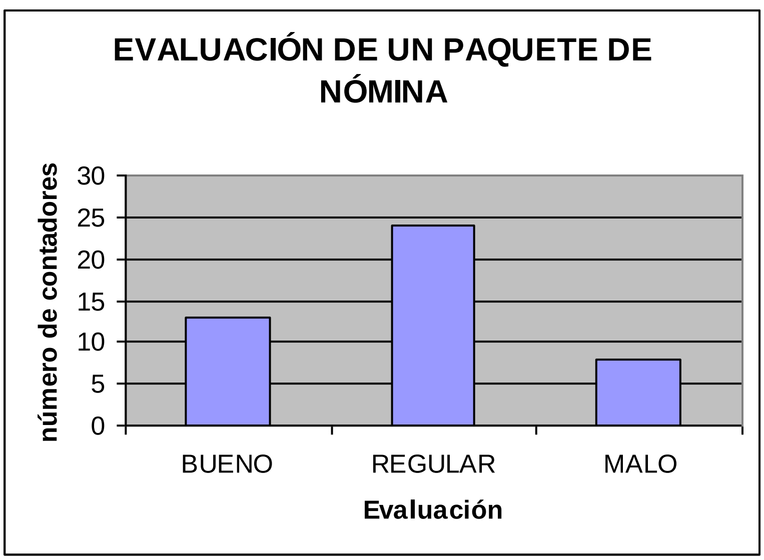
GRÁFICOS PARA DATOS CUALITATIVOS

Todo gráfico estadístico debe tener un título, relacionado con lo que se está investigando.

1. **DIAGRAMA O GRÁFICO DE BARRAS: puede ser de dos formas:**

2. **Diagrama o gráfico de barras verticales:** se hace en el primer cuadrante de un plano cartesiano que tendrá en el eje "X" las clases o categorías y en el eje "Y" las frecuencias absolutas.

Como información adicional en cada barra puede ir el porcentaje.

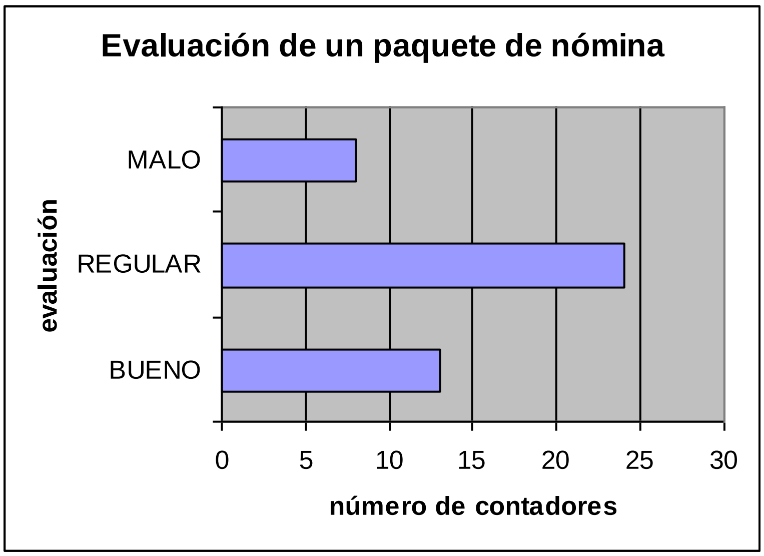


Interpretación del gráfico:
El grupo de contadores evaluó el paquete de nómina de la siguiente forma:
13 contadores lo evaluaron como bueno.

24 contadores lo evaluaron como regular.

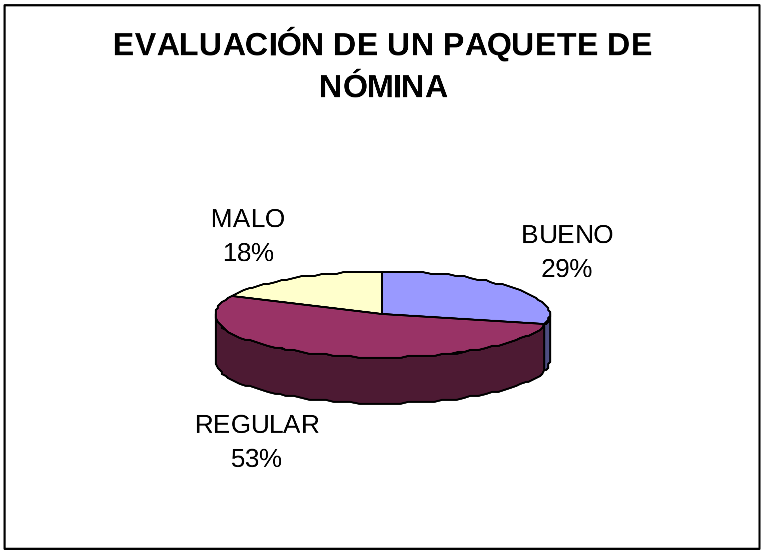
8 contadores lo evaluaron como malo.
Como podemos ver, esta interpretación es la misma que la frecuencia absoluta porque se refiere a cantidades.

3. **Diagrama o gráfico de barras horizontales**: en el eje "X" van las frecuencias y en el eje "Y" van las categorías



4. **Gráfico de sectores o diagrama circular:**

Es un diagrama que generalmente se utiliza para expresar los porcentajes, cada sector del gráfico indica la categoría y es proporcional a su porcentaje.



Siempre se debe seleccionar uno de los gráficos de acuerdo con lo que se quiere mostrar, cantidad o porcentaje.

Después de elaborar el gráfico, se realizan las conclusiones.

CONCLUSIONES:

- En general, el paquete no gustó entre los contadores, por lo tanto la empresa productora del software debe reformarlo basándose en las exigencias de los mismos.

Ejercicios tema 3.

TALLER SOBRE DATOS CUALITATIVOS

Para cada ejercicio:

- Elabore la ficha técnica.
- Elabore la tabla de distribución de frecuencias.
- Interprete la frecuencia absoluta y la relativa.
- Elabore el diagrama de barras y el circular.
- Exprese 2 conclusiones.

1. La compañía "los Dulces" desea lanzar al mercado una nueva chocolatina. Por tal razón realizó una encuesta entre los niños de las escuelas de Medellín para evaluar el producto y seleccionó una muestra al azar con los siguientes resultados (en miles):

Xi	**Fai**
Bueno	145
Regular	70
Malo	85

2. Entre las amas de casa de Medellín se elaboró una encuesta con el fin de investigar la efectividad de un detergente. Para tal fin se seleccionó una muestra al azar con los siguientes resultados (en miles):

Xi	**Fai**
Excelente	60
Bueno	145
Regular	20
Malo	25

3. Se realizó una encuesta entre los visitantes al Éxito de San Antonio en diciembre de 2002 para seleccionar entre 4 artículos: A1, A2, A3, A4 el de mejor calidad. se seleccionó una muestra al azar con los siguientes resultados (en miles):

Xi	Fai
A1	15
A2	80
A3	140
A4	9

4. El secretario de gobierno de Medellín ordenó un informe sobre las causas de muertes violentas ocurridas en Medellín durante el último trimestre, para tal fin seleccionó una muestra aleatoria y los resultados fueron los siguientes:

Xi	Fai
1. por accidentes de tránsito	45
2. por terrorismo	60
3. por arma de fuego	55
4. por otras causas.	15

5. Un investigador judicial realizó un informe sobre el número de reclusos que hay en las cárceles del Área Metropolitana discriminados por sexo, los resultados fueron los siguientes:

Xi	Fai
Masculino	7.586
Femenino	1.948

6. Para cada uno de los siguientes ejercicios, tome los datos de la siguiente tabla:

2	2	2	4	4	4	4	1	1	1	1	1	3	3	1
1	1	3	4	1	1	1	1	3	3	3	4	4	4	4
1	1	1	1	1	2	2	2	4	4	4	1	1	1	1
2	2	3	3	1	1	4	4	1	1	2	3	4	4	1

7. Un grupo de estudiantes de turismo de Corporación Universitaria Remington desean organizar un paseo entre los bachilleres de un colegio de Medellín. Para esto elaboraron un cuestionario donde el bachiller seleccionaba el sitio así: 1.Santa Fe de Antioquia, 2. La Pintada, 3. San Jerónimo, 4.El Peñol. Después de seleccionar una muestra al azar se tiene la siguiente información (ver datos de la tabla)

8. Un grupo de estudiantes de sistemas de Corporación Universitaria Remington elaboró 4 paquetes (software) sobre nómina. Después de hacer la demostración ante un grupo de gerentes, se les entregó un cuestionario para que eligieran el software más conveniente, así: 1.software1, 2.software2, 3.software3, 4.software4 (Ver datos de la tabla)

9. Los estudiantes de Mercadeo de Corporación Universitaria Remington desean montar una microempresa de dulces, para saber con cual producto van a iniciar, encuestaron a una muestra de los niños de un colegio de Medellín para que seleccionaran así: 1.bombón, 2.chocolatina, 3.bocadillo, 4.confites. (Ver datos de la tabla)

10. A los estudiantes del último semestre de Contaduría Pública de Corporación Universitaria Remington se les preguntó en que les gustaría desempeñarse cuando fueran profesionales: 1. contador general, 2. docente, 3.auditor, 4.revisor fiscal. (Ver datos de la tabla)

11. En la institución “Despertar” el psicólogo clasificó a los niños sobre el grado de retardo mental según el DSMIV: 1. leve, 2. moderado, 3. grave 4. Profundo. (Ver datos de la tabla)

12. En un cuestionario que el jefe de personal de la empresa “Bienestar” le hizo a sus empleados, una de las preguntas está diseñada así: su estado civil es: 1.soltero, 2.casado, 3.separado, 4.unión libre, 5.viudo. Al recolectar los datos se tienen los siguientes resultados:

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

13. El sociólogo de la empresa “Unidas” realizó un estudio entre los empleados, escogiendo algunos aleatoriamente. Una de las preguntas era sobre el tipo de vivienda, la información recolectada fue la siguiente: 1.propia sin deuda, 2. propia con deuda, 3.arrendada, 4.prestada

3	3	3	3	1	1	2	2	2	3
3	3	4	4	3	3	2	2	2	1
1	3	3	3	2	2	2	4	4	2
2	2	3	3	3	3	2	2	1	1
4	4	3	3	3	1	3	3	3	3

14. Un grupo de contadores públicos y de ingenieros de sistemas de Corporación Universitaria Remington elaboró un software contable; se eligió una muestra aleatoria entre los contadores públicos de Medellín a los cuales le entregaron el paquete y lo evaluaron así:

Xi	fai
Excelente	500
Bueno	800
Regular	50
Deficiente	100

15. A los extranjeros que visitaron la embajada de Colombia en Estados Unidos durante el último trimestre se les preguntó sobre su visita a Colombia y respondieron lo siguiente:

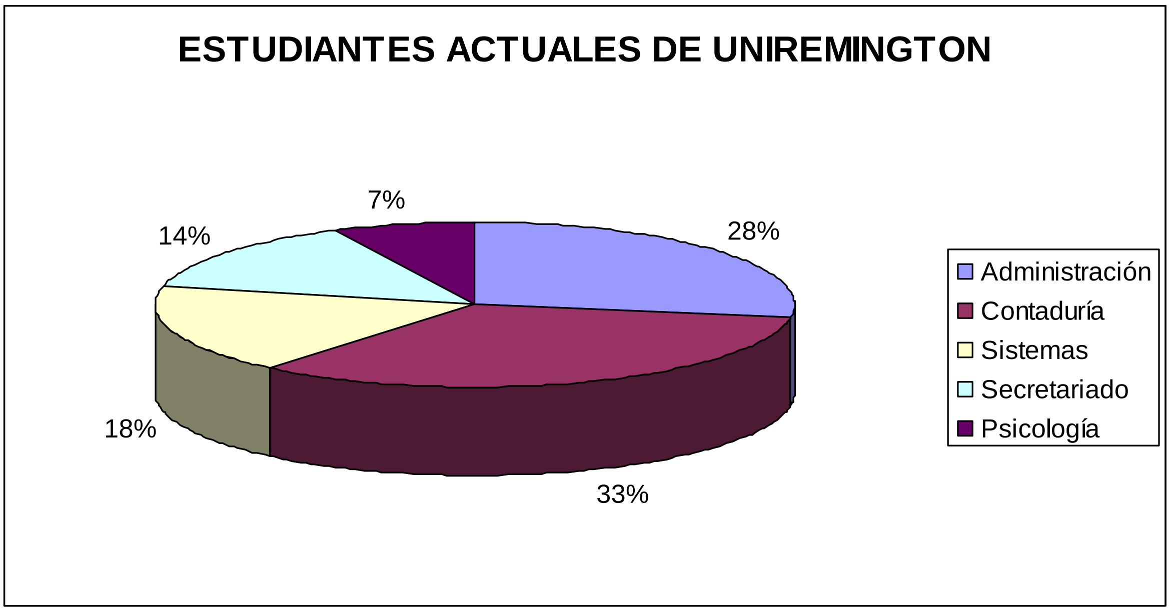
Xi	fai
Vendrán	167
Nunca vendrán	16
Ya vinieron y volverán	70
Ya vinieron y nunca volverán	6

16. Se hizo una selección aleatoria entre los estudiantes de Corporación Universitaria Remington y se les preguntó ¿En qué dedican su tiempo libre? y respondieron lo siguiente:

Xi fai

Ir a cine	178
Hacer deporte	150
Ir a discoteca	200
Leer	40

Actividad final de la unidad 1



Responda falso o verdadero:

1. Esta investigación se hizo con datos cuantitativos.
2. La mayoría de los estudiantes son los de Contaduría.
3. Se debe promocionar más la Psicología.
4. Lo que menos se presenta es que haya estudiantes de Secretariado.
5. Este gráfico es el único que se puede elaborar para esta investigación.
6. La población de esta investigación son los estudiantes de Corporación Universitaria Remington.

Actividad: Realiza una investigación estadística, sobre datos cualitativos en tu medio; ya sea tu lugar de trabajo, tu ciudad o tu familia y realiza todo el proceso: tablas, gráficos y conclusiones o decisiones finales.

4. DATOS CUANTITATIVOS ORDENADOS EN FILA

OBJETIVO GENERAL

Describir por medio de tablas, gráficos y medidas los datos ordenados en fila.

OBJETIVOS ESPECÍFICOS

- Interpretar los datos: mayor, menor, más frecuente y menos frecuente en un conjunto de datos cuantitativos ordenados en fila.
- Construir las frecuencias acumuladas e interpretarlas.
- Realizar el diagrama de barras e interpretarlo.
- Calcular las medidas de tendencia central e interpretarlas.

Prueba inicial

Se seleccionó una muestra aleatoria entre los estudiantes de Corporación Universitaria Remington y clasificaron de acuerdo con su edad, los siguientes son los datos:

Xi(edad)	fai(nº est)	fri	%	faai	frai	Xi*fai
17	100					
19	140					
23	100					
26	140					
29	100					
35	25					
45	8					

Termine la tabla y responda a las siguientes preguntas, sustentando su respuesta:

1. El dato más frecuente es 100 (F) (V)
2. Lo que menos se presenta es que los estudiantes tengan 35 años (F) (V)
3. Hay 8 estudiantes que tienen la edad mayor (F) (V)

4 .El 27% de los estudiantes tiene menos de 26 años (F) (V)
5. Hay 140 estudiantes que tienen como máximo 26 años (F) (V)
6. Hay 340 estudiantes que tienen como máximo 23 años (F) (V)
7. El promedio de la edad entre los estudiantes es de: a) 43,545 b) 25,758 c) 23,548
8. Las edades más frecuentes son: a) 100 y 140 b) 35 y 25 c) 17 y 19 d) 19 y 26
9. El 50% de las edades entre los estudiantes es de: a) 23 b) 19 c) 26
10. El 73,5% de los estudiantes tienen 29 años o menos (F) (V)

4.1. Tablas para datos cuantitativos ordenados en fila

DATOS CUANTITATIVOS

Cuando la investigación es de tipo cuantitativa, es decir que se tienen datos numéricos, los datos se pueden organizar de dos formas: ordenados en una fila o agrupados en intervalos de clase.

DATOS ORDENADOS EN FILA

Cuando los datos numéricos son pocos, o siendo muchos, cada uno representa una alta frecuencia. Los datos se organizan en una fila en forma ascendente.

Ejemplo:
El jefe de personal de la compañía "Aceros S.A." preocupado por las llegadas tarde de sus empleados, seleccionó una muestra aleatoria entre los empleados que han llegado tarde al trabajo durante los últimos cuatro meses y anotó el número de llegadas tarde de cada uno de dichos empleados y los resultados fueron los siguientes:

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

FICHA TÉCNICA
Población: Empleados que han llegado tarde al trabajo durante los últimos cuatro meses de la compañía "Aceros S.A."
Muestra: 60 empleados escogidos al azar.
Descripción de la variable: Número de llegadas tarde de cada empleado.
Tipo de variable: Cuantitativa discreta.

TABLA DE DISTRIBUCIÓN DE FRECUENCIAS

Xi	fai	fri	%	faai	frai	% Acum
1	5	0,083	8,3	5	0,083	8,3
2	4	0,067	6,7	9	0,150	15,0
3	8	0,133	13,3	17	0,283	28,3
4	8	0,133	13,3	25	0,417	41,7
5	21	0,350	35,0	46	0,767	76,7
6	8	0,133	13,3	54	0,900	90,0
7	2	0,033	3,3	56	0,933	93,3
8	4 n = 60	0,067 ∑ = 1	6,7 ∑ = 100	60	1,000	100,0

Interpretación de los datos: mayor, menor, más frecuente y menos frecuente

Como los datos fueron organizados de menor a mayor, entonces siempre el primer dato será el menor y el último dato será el mayor; además recuerde que los datos siempre están en **Xi** y que la frecuencia absoluta indica el número de veces que se presenta el dato.

1. **DATO MENOR:** El menor número de llegadas tarde de los empleados es de **1** y hubo 5 empleados que representan el 8,3% que llegaron tarde al trabajo 1 vez.

2. **DATO MAYOR:** El mayor número de llegadas tarde de los empleados es de **8** y hubo 4 empleados que representan el 6,7% que llegaron tarde al trabajo 8 veces.

3. **DATO MÁS FRECUENTE:** Lo que más se presenta es que los empleados lleguen tarde al trabajo **5 veces;** 21 empleados que representan el 35% llegaron tarde 5 veces.

4. **DATO MENOS FRECUENTE:** Lo que menos se presenta es que los empleados lleguen tarde al trabajo 7 veces; 2 empleados que representan el 3,3% llegaron tarde 7 veces.

A continuación se explican otros datos en la tabla:

Frecuencia absoluta acumulada (faai o Ni): Es sumar las frecuencias absolutas por cada categoría: faai = fai + $faa_{(i-1)}$
La última frecuencia absoluta acumulada es igual a N ó n.

Frecuencia relativa acumulada (frai o Hi): Es sumar las frecuencias relativas por cada categoría: frai = fri + $fra_{(i-1)}$ o también $frai = \frac{faai}{N \text{ ó } n}$

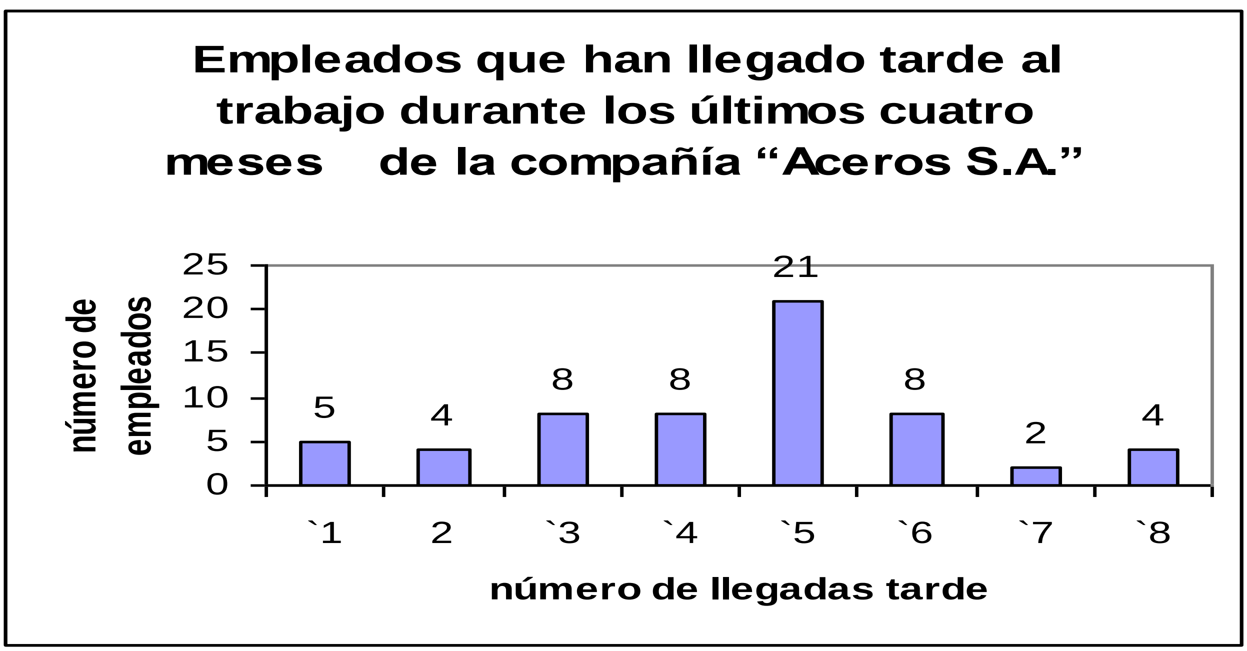
La última frecuencia relativa acumulada es igual a uno.

Porcentaje acumulado: Se obtiene multiplicando las frecuencias relativas acumuladas por cien.

GRÁFICOS PARA DATOS ORDENADOS EN FILA

Se utilizan los mismos diagramas de los datos cualitativos, aunque el diagrama circular no es muy usual porque se presta a confusiones pues queda el valor y el porcentaje en las categorías.

A continuación se tiene el diagrama de barras verticales de la anterior investigación.



Ejercicios tema 1

En cada uno de los siguientes ejercicios, realice lo siguiente:

- Ficha técnica
- Interprete dato: mayor, menor, más frecuente y menos frecuente.
- Elabore el diagrama de barras e interprételo.

1. La profesora de estadística realizó una prueba a un grupo de alumnos que constaba de 12 puntos. Eligió una muestra al azar y los resultados del número de respuestas acertadas fueron los siguientes:

12	10	10	8	4	3	0	2	1	12
12	11	9	2	6	4	8	0	7	11
5	9	9	3	7	6	9	0	8	8
4	10	9	7	8	7	9	1	7	8
6	10	4	8	11	10	7	12	9	8

2. En la feria Exponavidad del año pasado se realizó un estudio sobre lo que los visitantes tenían disponible para gastar en regalos navideños. Se seleccionó una muestra al azar con los siguientes resultados: (el dinero disponible está dado en miles)

Xi	**fai**
50	10
200	25
400	15
600	10
800	8
1000	8
1500	2
2000	3

4.2. Medidas de tendencia central para datos ordenados en fila

Son medidas que centralizan los datos y dan información sobre la parte de la distribución hacia donde se están agrupando los datos. Las más importantes son: Media aritmética, la moda y la mediana.

MEDIA ARITMÉTICA

"Es la medida más conocida, más fácil de calcular y con la que siempre estamos más familiarizados" (Martínez Bencardino, C. 2004, p. 74) Es el promedio de los datos y se representa:
µ para un censo es decir, como parámetro
X para un muestreo es decir, como estadística.

Para calcular la media aritmética: se debe tener en cuenta lo siguiente:

1. **si los datos no están ordenados:** "Se obtiene sumando todos los valores y dividiendo por la cantidad de valores" (Anderson, D., Sweeney, D. y Williams, T., 1999, p.65).

$$\mu \text{ ó } X = \frac{x_1 + x_2 + \ldots + x_n}{N \text{ ó } n}$$

Ejemplo 1: a continuación se tienen las notas de seguimiento en estadística de un estudiante 4,5 3,8 4,5 4,0 4,5

$$\mu = \frac{4{,}5 + 3{,}8 + 4{,}5 + 4 + 4{,}5}{5} = 4{,}26$$

El promedio de las notas de estadística del alumno es de 4,26

Ejemplo 2: a continuación se dan los sueldos de los empleados del almacén "El Reloj" de acuerdo con una muestra aleatoria: 590, 590, 880, 1000, 590, 590, 1000 (en miles)

$$\dddot{X} = \frac{590 + 590 + 880 + 1000 + 590 + 590 + 1000}{7} = 748{,}571$$

El promedio de los sueldos de los empleados del almacén el administrador es de $748571

2. cuando se tienen ordenados en una tabla de distribución de frecuencias para hallar la media aritmética se aplica la siguiente fórmula:

$$\mu \text{ ó } \overline{x} = \frac{x_{1.} fa_1 + x_2 . fa_2 + ... + x_n . fa_n}{N \text{ ó } n}$$

Tomando el ejemplo de la tabla de frecuencias:

Xi	**fai**	**Xi*fai**
1	5	5
2	4	8
3	8	24
4	8	32
5	21	105
6	8	48
7	2	14
8	4	32
	n = 60	∑ = 268

$$\overline{X} = \frac{268}{60} = 4{,}467$$

El promedio de las llegadas tarde de los empleados de la compañía "Aceros S.A." es de 4,467 (si se analiza, llevándolo a la realidad, es que en promedio llegan 4 veces tarde al trabajo)
Propiedades de la media aritmética:

1. Todo conjunto de datos cuantitativos tiene una media aritmética.
2. Cuando se calcula la media aritmética se incluyen todos los datos.
3. La media aritmética de un conjunto de datos es única es decir, sólo existe una media aritmética.
4. Cuando se tienen valores extremos (es decir, muy altos o muy bajos) la media aritmética no puede tomarse como representativa de los datos.

LA MODA (Mo)

"Es el valor que ocurre con mayor frecuencia; es decir, el valor más frecuente" (Spiegel, M. R., 1995).

Cuando se tiene la distribución de frecuencias se busca la mayor frecuencia absoluta y su correspondiente "Xi" es la moda.

Tomando los ejemplos anteriores, la moda se interpreta así:

Ejemplo 1: La nota de seguimiento en estadística del estudiante más frecuente es 4,5

Ejemplo 2: El sueldo de los empleados del almacén “El Reloj” más frecuente es $ 590000

Ejemplo 3: Las llegadas tarde de los empleados de la compañía “Aceros S.A.” Más frecuente es de 5

Existen distribuciones que tienen varias modas por lo tanto se denomina **bimodal** cuando tiene dos modas, si tiene más de dos, se llama **multimodal**.

Pueden existir distribuciones donde no haya moda, si todos los valores de la frecuencia absoluta son iguales, no hay moda, pues indica que todos los datos se repiten el mismo número de veces.
En el gráfico de barras verticales la moda se identifica como barra más alta y su valor se ubica en el eje X

LA MEDIANA (Me)
“Es el valor de la serie de datos que se sitúa justamente en el centro de la muestra (un 50% de valores son inferiores y otro 50% son superiores).
No presentan el problema de estar influido por los valores extremos, pero en cambio no utiliza en su cálculo toda la información de la serie de datos (no pondera cada valor por el número de veces que se ha repetido).” (http://www.aulafacil.com/CursoEstadistica/CursoEstadistica.htm)

Es el valor que está en todo el centro de la distribución. Para ubicarla los datos deben estar ordenados. Tomando los ejemplos anteriores, la mediana se interpreta así:

Ejemplo 1:
3,8 4,0 4,5 4,5 4,5
El 50% de las notas de seguimiento en estadística del estudiante es 4,5

Ejemplo 2:
590, 590, 590, 590, 880, 1000, 1000

El 50% de los sueldos de los empleados del almacén “El Reloj” es $590000

Cuando se tiene un número de datos par se toman los dos valores centrales y se promedian.

Ejemplo:

3,8 3,8 4,0 4,5 4,5 4,5 $Me = \frac{40 + 4{,}5}{2} = 4{,}3$

Cuando se tiene la distribución de frecuencias para ubicar la mediana se procede así:

1. **Si el número de datos (N ó n) es impar** se calcula $\frac{n}{2}$, se busca en la faa (frecuencia absoluta acumulada) y su correspondiente dato (Xi) es la Me.

2. **Si el número de datos (N ó n) es impar** se calcula $\frac{n+1}{2}$, se busca en la faa (frecuencia absoluta acumulada) y su correspondiente dato (Xi) es la Me.

xi	**fai**	**faai**
3.8	1	1
4.0	1	2
4.5	3	5
	N=5	

$Me = \frac{5+1}{2} = 3$ Corresponde al tercer dato que es 4.5

3. Si el número de datos (N o n) es par se calcula $\frac{n}{2}$ y su consecutivo, se buscan en la faa (frecuencia absoluta acumulada) y el promedio de sus correspondientes datos (Xi) es la Me.

Xi	**fai**	**faai**
3,8	2	2
4,0	1	3
4,5	3	6
	N = 6	

$Me = \frac{6}{2} = 3$ Y su consecutivo es 4 $Me = \frac{4{,}0 + 4{,}5}{2} = 4{,}3$

En el ejemplo de la compañía “Aceros S.A.” N/2 = 60/2 = 30 y su consecutivo 31 como ambos están incluidos en 46, la Mediana sería 5

El 50% de las llegadas tarde de los empleados de la compañía “Aceros S.A.” es de 5

Ejercicios tema 2

1. La profesora de estadística realizó una prueba a un grupo de alumnos que constaba de 12 puntos. Eligió una muestra al azar y los resultados del número de respuestas acertadas fueron los siguientes:

12	10	10	8	4	3	0	2	1	12
12	11	9	2	6	4	8	0	7	11
5	9	9	3	7	6	9	0	8	8
4	10	9	7	8	7	9	1	7	8
6	10	4	8	11	10	7	12	9	8

2. En la feria Exponavidad del año pasado se realizó un estudio sobre lo que los visitantes tenían disponible para gastar en regalos navideños. Se seleccionó una muestra al azar con los siguientes resultados: (el dinero disponible está dado en miles)

Xi	**fai**
50	10
200	25
400	15
600	10
800	8
1000	8
1500	2
2000	3

Calcule las medidas de tendencia central de los anteriores ejercicios.

Actividad final de la unidad 2

Se seleccionó una muestra aleatoria entre los habitantes de la tercera edad que viven en el Área Metropolitana y se clasificaron de acuerdo con su estatura, los siguientes son los datos:

Xi(estat)	fai(nºancia)	fri	%	faai	frai	Xi*fai
1,45	6					
1,50	10					
1,55	10					
1,60	18					
1,65	18					
1,70	12					
1,75	10					
1,80	7					
1,85	3					
1,90	1					

Termine la tabla y responda a las siguientes preguntas, sustentando su respuesta:

1 El dato más frecuente es 10 (F) (V)
2 Lo que menos se presenta es que los ancianos midan 1,45 (F) (V)
3 Esta distribución es Bimodal (F) (V)
4 Lo que más se presenta es que los ancianos midan 1,90 (F) (V)
5 Hay 26 ancianos que miden menos de 1,60 (F) (V)
6 El 58% de los ancianos miden como máximo 1,80 (F) (V)
7 El promedio de la estatura entre los ancianos es de: a) 1,55 b) 1,60 c) 1,638
8 La estatura más frecuente es: a) 10 b) 1,90 c) 1,60 y 1,65 d) 1,50 1,55 y 1,75
9 El 50% de las estaturas entre los ancianos es de: a) 1,60 b) 1,70 c) 1,65
10 El 28% de los ancianos miden 1,65 o menos (F) (V)

Actividad:
Realiza una investigación estadística, sobre datos cuantitativos en fila en tu medio; ya sea tu lugar de trabajo, tu ciudad o tu familia y realiza todo el proceso: tablas, gráficos, medidas y conclusiones o decisiones finales.

5. DATOS CUANTITATIVOS AGRUPADOS EN INTERVALOS DE CLASE

OBJETIVO GENERAL

Analizar datos cuantitativos agrupados en intervalos de clase.

OBJETIVOS ESPECÍFICOS

- Organizar una serie de datos numéricos en intervalos.
- Construir e interpretar las frecuencias acumuladas.
- Construir gráficos para datos agrupados en intervalos.
- Calcular, interpretando los resultados, las medidas de tendencia central para datos agrupados en intervalos.
- Calcular e interpretar las medidas de posición relativa para datos agrupados en intervalos.
- Calcular e interpretar las medidas de variación o dispersión para datos cuantitativos.

Prueba Inicial

En la compañía "Universal S. A." se hizo un estudio sobre los sueldos, a continuación se dan los resultados, en miles de $.

TABLA DE DISTRIBUCIÓN DE FRECUENCIAS									
LI	LS	fai	fri	%	faai	frai	xi	xi . Fai	$(xi - \mu)^2$. Fai
297	748	15	0.375	37.5	15	0.375	522.5	7837.5	4578429.38
748	1199	10	0.25	25	25	0.625	973.5	9735	102971.756
1199	1650	9	0.225	22.5	34	0.85	1424.5	12820.5	1099509.53
1650	2101	4	0.1	10	38	0.95	1875.5	7502	2563361.1
2101	2552	1	0.025	2.5	39	0.975	2326.5	2326.5	1566314.83
2552	3003	1	0.025	2.5	40	1	2777.5	2777.5	2898591.38
		40	1	100				42999	12809178

De acuerdo con los datos de la tabla, responde lo siguiente:

El promedio de sueldos de la compañía "Universal S.A" es de:

a) $ 1.074.975 **b)** $ 2.650.000 **c) $** 1.5689.532 **d) $** 3.420.000

El 50%: de sueldos de la compañía "Universal S.A" es de:

a) $ 973.500 **b)** $ 2.650.000 **c) $** 1.5689.532 **d) $** 3.420.000

El sueldo de la compañía "Universal S.A" más frecuente es:

a) $ 973.500 **b)** $ 2.650.000 **c) $** 1.5689.532 **d)** $ 635.250.

Se puede afirmar que en la compañía "Universal S.A" la mayoría de los sueldos son:

a) Altos **b)** Regulares **c)** Bajos

5.1. Tablas para Datos cuantitativos agrupados en intervalos de clase

Cuando se tienen muchos datos y cada uno tiene una frecuencia baja, es decir que no se repiten o se repiten pocas veces es necesario agrupar estos datos; en cada clase o categoría de la distribución irá un intervalo. Existen vario métodos para agruparlos; uno de ellos es que los intervalos sean de la forma: **LI < X ≤ LS** o sea que se incluye el LS (límite superior) y no se incluye LI (límite inferior) se deben seguir los siguientes pasos:

1**. Se calcula el rango**

R = dato mayor – dato menor

2**. Se calcula el # de intervalos (k) que tendrá la distribución**

K = 1 + 3,3*log n donde n es el número de datos

K se redondea a entero

3**. Se calcula la amplitud que va a ser la misma en todos los intervalos**

A = R/K

La amplitud se aproxima al entero siguiente así sea entero o sea se le suma 1 a la amplitud.

4**. Se calcula el rango ampliado**

Ra = A*K

5**. Se calcula una cantidad que debe ser diferente de cero**

$$\text{Cant} = \frac{Ra - R}{2}$$

6**. Se calcula el primer límite inferior que tendrá la distribución.**

LI = dato menor – cant

Ejemplo 1:
Entre los estudiantes de Administración de Corporación Universitaria Remington se hizo una selección aleatoria y se analizó la edad en que iniciaron sus estudios en la universidad, los siguientes son los datos:

25	36	40	45	48	32	17	22	18	20
21	44	46	19	19	26	24	28	29	31
39	36	35	33	32	34	41	42	31	30
20	45	43	23	22	21	38	34	37	43
19	20	23	25	22	25	27	24	19	18
20	21	22	25	24	27	29	19	23	21

Ficha técnica:
POBLACIÓN: Estudiantes de Administración de Corporación Universitaria Remington.
MUESTRA: 60 estudiantes elegidos al azar.
DESCRIPCIÓN DE LA VARIABLE: La edad en que iniciaron sus estudios en la universidad.
TIPO DE VARIABLE: Cuantitativa continúa (debido a que pertenece a un intervalo)

Como podemos apreciar, existen muchos datos con baja frecuencia, por tanto hay que agruparlos en intervalos; vamos a aplicar los pasos:

1**. Se calcula el rango**
R = dato mayor – dato menor
R = 48 – 17 = 31

2. **Se calcula el # de intervalos (k) que tendrá la distribución**
K = 1 + 3,3*log n donde n es el número de datos
K = 1 + 3,3 *log 60 = 6,868
K se redondea a entero K = 7

3**. Se calcula la amplitud que va a ser la misma en todos los intervalos**
A = R/K A = 31/7 = 4,428
La amplitud se aproxima al entero siguiente así sea entero o sea se le suma 1 a la amplitud. A = 5

4**. Se calcula el rango ampliado**
Ra = A*K = 5*7 = 35

5**. Se calcula una cantidad que debe ser diferente de cero**
Cant = Ra – R = 35 -31 = 4

6**. Se calcula el primer límite inferior que tendrá la distribución.**
LI = dato menor – cant
LI = 17 – 2 = 15

Como todos los intervalos tiene la misma amplitud, **LS = LI + A**
Primer límite superior: LS = 15 + 5 = 20

Recordemos que vamos a agrupar intervalos de la forma **LI < X ≤ LS** para el primer intervalo contamos los datos que son mayores que 15 y menores o iguales que 20 y en total encontramos 12 datos esta será la primera frecuencia absoluta.

Todo límite superior pasa a ser límite inferior en el siguiente intervalo y para calcular el segundo límite superior aplicamos la fórmula **LS = LI + A** por tanto el segundo intervalo tendrá como límite inferior 20 y como límite superior 20 + 5 = 25 y contamos los datos que son mayores que 20 y menores o iguales que 25 y en total encontramos 12 datos esta será la segunda frecuencia absoluta. Así sucesivamente.

TABLA DE DISTRIBUCIÓN DE FRECUENCIAS

LI	LS	fai	fri	%	Xi	faai	frai	% acum
15	20	12	0,2	20	17,5	12	0,2	20
20	25	18	0,3	30	22,5	30	0,50	50
25	30	7	0,117	11,7	27,5	37	0,617	61,7
30	35	8	0,133	13,3	32,5	45	0,75	75
35	40	6	0,1	10	37,5	51	0,85	85
40	45	7	0,117	11,7	42,5	58	0,967	96,7
45	50	2	0,33	3,3	47,5	60	1	100
		N = 60	∑ = 1	∑ = 100				

Ejemplo 2:
Los siguientes son los sueldos de los tecnólogos de Medellín de acuerdo con una muestra elegida aleatoriamente (en miles de pesos).
En este caso ya se dan los datos agrupados, por tanto no hay que aplicar los pasos, simplemente se complementa la tabla.

LI	LS	fa
400	600	30
600	800	100
800	1000	55
1000	1200	240
1200	1400	25
1400	1600	20

TABLA DE DISTRIBUCIÓN DE FRECUENCIAS

LI	LS	fai	fri	%	Xi	faai	frai	% acum
400	600	30	0,064	6,4	500	30	0,064	6,4
600	800	100	0,213	21,3	700	130	0,277	27,7
800	1000	55	0,117	11,7	900	185	0,394	39,4
1000	1200	240	0,511	51,1	1100	425	0,904	90,4
1200	1400	25	0,053	5,3	1300	450	0,957	95,7
1400	1600	20	0,043	4,3	1500	470	1,000	100,0
		n = 470	∑ = 1	∑ = 100				

GRÁFICOS PARA DATOS AGRUPADOS

1. **EL HISTOGRAMA**: "Son diagramas de barras verticales en los que se construyen barras rectangulares en los límites de cada clase." (Berenson, M. L. Y LEVINE, D. M. 1996, p. 70).

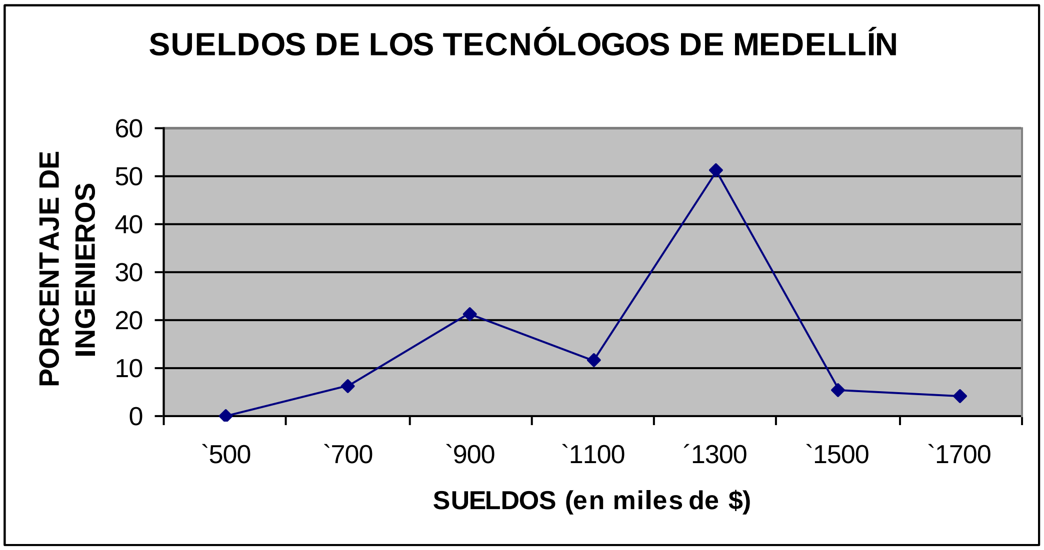
Puede ser de frecuencias absolutas o de porcentajes; se realiza colocando en el eje X los límites y en el eje Y las frecuencias absolutas o los porcentajes Según el caso.

Marcas de clase (Xi):
Es un valor que identifica a cada intervalo "es el punto medio de cada intervalo" (Berenson, M. L. Y LEVINE, D. M., 1996, p. 38).

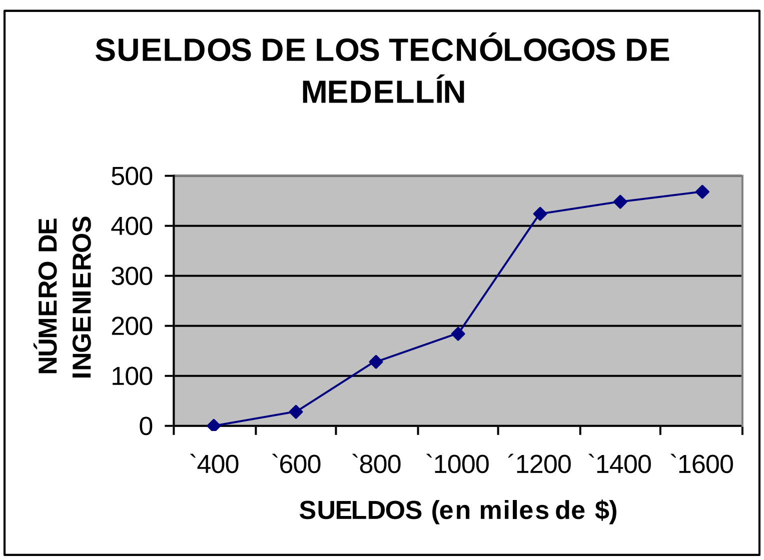
$$xi = \frac{LI + LS}{2}$$

2. **POLÍGONO: "**Se puede obtener uniendo cada punto medio (marca de clase) de los rectángulos del histograma con líneas rectas, teniendo cuidado de agregar al inicio y al final marcas de clase adicionales, con el objeto de asegurar la igualdad del áreas". (http://sitios.ingenieria-usac.edu.gt/estadistica/estadistica2/estadisticadescriptiva.html)

Puede ser de frecuencias absolutas o de porcentaje. Se realiza colocando en el eje X las marcas de clase y en el eje y las frecuencias absolutas o los porcentajes según el caso. El siguiente es el polígono de porcentaje del ejemplo 2



3. **OJIVA:** puede ser de frecuencias absolutas acumuladas o de porcentajes acumulado. Se realiza colocando en el eje X los límites y en el eje Y las frecuencias acumuladas. La siguiente ojiva es de faai del ejemplo 2



Ejercicios tema 1

En cada uno de los siguientes ejercicios realice:

- La ficha técnica.
- Interprete el intervalo: mayor, menor, más frecuente y menos frecuente.
- Elabore el Histograma, el Polígono y la Ojiva.
- Calcule e interprete: las medidas de tendencia central, las medidas de posición relativas y las medidas de variabilidad.

1. En la compañía "La Delicia" se hizo un estudio sobre los sueldos, a continuación se dan los resultados, en miles de $.

900	500	450	1900	1200	1250	2500	550	1650	1200
1000	550	950	600	750	1300	850	350	1400	700
300	1100	300	600	1600	1500	1000	1800	900	500
650	2000	450	750	850	600	300	1950	3000	1500

5.2. Medidas de tendencia central para datos cuantitativos agrupados en intervalos de clase

1. Media aritmética:

$$\overline{X} \text{ ó } \mu = \frac{\sum Xi.fai}{n \text{ ó } N}$$

Para el ejemplo 1

LI	**LS**	**fai**	**Xi**	**Xi*fai**
15	20	12	17,5	210
20	25	18	22,5	405
25	30	7	27,5	192,5
30	35	8	32,5	260
35	40	6	37,5	225
40	45	7	42,5	297,5
45	50	2	47,5	95
		N = 60		∑ = 1685

$$\bar{X} = \frac{1685}{60} = 28{,}03$$

El promedio de la edad en que iniciaron sus estudios en la universidad los estudiantes de Administración de Corporación Universitaria Remington es de 28,03 años

Para el ejemplo 2

LI	LS	fai	Xi	Xi*fai
400	600	30	500	15000
600	800	100	700	70000
800	1000	55	900	49500
1000	1200	240	1100	264000
1200	1400	25	1300	32500
1400	1600	20	1500	30000
		n = 470		∑ = 461000

$$\bar{X} = \frac{461000}{470} = 980{,}851$$

El promedio de los sueldos de los tecnólogos de Medellín es de $980.851

4. **Moda:**

$$Mo = LI + \left(\frac{d1}{d1 + d2}\right) * A$$
d1 = fai – fa(i-1) d2 = fai – fa(i+1) A = LSi – LIi

El intervalo que contiene la Moda es el más frecuente
El intervalo más frecuente es el **rojo** en las tablas, por lo tanto:

Para el ejemplo 1

LI = 20 d1 = 18 -12 = 6 d2 = 18 – 7 = 11 A = 25 – 20 = 5 reemplazando

$$Mo = 20 + \left(\frac{6}{6 + 11}\right) * 5 = 21{,}764$$

La edad en que iniciaron sus estudios en la universidad los estudiantes de Administración de Corporación Universitaria Remington más frecuente es de 21,764 años

Para el ejemplo 2

LI = 1000 d1 = 240 -55 = 185 d2 = 240 – 25 = 215 A = 1200 – 1000 = 200

Reemplazando en la fórmula:

$$Mo = 1000 + (\frac{185}{185+215})*200 = 1092{,}5$$

Como está dado en miles, se multiplica 1092,5*1000 = 1`092.500

El sueldo de los tecnólogos de Medellín más frecuente es de $1`092.500

5**. Mediana:**

$$Me = LI + (\frac{n/2 - faa(i-1)}{fai})*A$$

Para ubicar el intervalo que contiene la mediana se puede hacer por:

1. n/2 y se busca en faa

2. 50 % y se busca en porcentaje acumulado

Para el ejemplo 1

El intervalo que contiene la mediana lo ubicamos n/2 = 60/2 = 30 este valor, buscando en la frecuencia absoluta acumulada, está en el 2º intervalo es el mismo de la moda, pero no siempre dan en el mismo, existen ciertas distribuciones que cumplen con esta característica. Si observamos, acá está el 50% en el porcentaje acumulado. Reemplazando en la fórmula:

$$Me = 20 + (\frac{30-12}{18})*5 = 25$$

El 50% de la edad en que iniciaron sus estudios en la universidad los estudiantes de Administración de Corporación Universitaria Remington es de 25 años.

Para el ejemplo 2

El intervalo que contiene la mediana lo ubicamos n/2 = 235 este valor está en 425 en el mismo de la moda, no siempre dan en el mismo, existen ciertas distribuciones que cumplen con esta característica. Si observamos, acá está el 50% en el porcentaje acumulado. Reemplazando en la fórmula:

$$Me = 100 + (\frac{235-185}{240})*200 = 1041{,}667$$

Como está dado en miles, se multiplica 1041,667*1000 = 1`041.667

El 50% de los sueldos de los tecnólogos de Medellín es de $1`041.667

Ejercicios tema 2

1**.** A continuación se dan las notas de los estudiantes de sistemas de acuerdo con una muestra elegida al azar:

LI	**LS**	**fai**
0	1	15
1	2	15
2	3	20
3	4	28
4	5	22

Calcule e interprete: La Media Aritmética, La Moda y la Mediana de ambos ejercicios.
Calcule e interprete: para el ejercicio 1 Q1, D3, P5 y para el ejercicio 2, Q3, D9 y P55

5.3. Medidas de posición relativa para datos cuantitativos agrupados en intervalos de clase

Estas medidas dividen la distribución en partes iguales, así como La Mediana, por tanto se calculan e interpretan similar a ella. Se tienen las siguientes:

1**. Cuartiles (Q):** dividen la distribución en 4 partes iguales así:
Q1=25%
Q2=50%
Q3=75%
Q4=100%

$$Q = LI + (\frac{{}^{\#Q*n}\!/_{4} - faa(i-1)}{fai})*A$$

Para ubicar el intervalo que contiene Q se puede hacer por

1) $\frac{\#Q*n}{4}$ y se busca en faa

2) el porcentaje respectivo se busca en el porcentaje acumulado

Si nos piden el cuartil 1 del ejemplo 2:

Q_1 = 1*470/4 = 117,5 en la faai está en 130 o el 25% está acá también, en la tabla es el intervalo azul. Reemplazando en la fórmula:

$$Q_1 = 600 + (\frac{117{,}5 - 30)}{100})*200 = 775$$

El 25% de los sueldos de los tecnólogos de Medellín es de $ 775.000 o menos.

2. **Deciles (D):** “dividen la distribución en 10 partes iguales” (Murray, 1995, p. 66) así:
D1=10%
D2=20%

D10=100%

$$D = LI + (\frac{{}^{\#D*n}\!/_{10} - faa(i-1)}{fai})*A$$

Para ubicar el intervalo que contiene D se puede hacer por

1) $\frac{\#D*n}{10}$ y se busca en faa

2) el porcentaje respectivo se busca en el porcentaje acumulado

Si nos piden el decil 3 del ejemplo 2:
D_3 = 3*470/10 = 141 en la faai está en 185 o el 30% está acá también, en la tabla es el intervalo **marrón**. Reemplazando en la fórmula:

$$D_3 = 800 + (\frac{141-130)}{55}) * 200 = 840$$

El 30% de los sueldos de los tecnólogos de Medellín es de $ 840.000 o menos.

3**. Percentiles (P):** “dividen la distribución en 100 partes iguales” (Murray, 1995, p. 66), así:
P1=1%
P2=2%

P100=100%

$$P = LI + (\frac{\#P*n/_{100} - faa(i-1)}{fai}) * A$$

Para ubicar el intervalo que contiene P se puede hacer por
1) $\frac{\#P*n}{100}$ y se busca en faa

2) el porcentaje respectivo se busca en el porcentaje acumulado.
Si nos piden el percentil 5 del ejemplo 2:
P_5 = 5*470/100 = 23,5 en la faai está en 30 o el 5% está acá también, en la tabla es el intervalo fucsia, el primero. Reemplazando en la fórmula:

$$P_5 = 400 + (\frac{23,5-0)}{30}) * 200 = 556,667$$

El 5% de los sueldos de los tecnólogos de Medellín es de $ 556.667 o menos.

Ejercicios tema 3

Para cada uno de los siguientes ejercicios calcule e interprete las medidas de variabilidad y tome sus propias conclusiones.

1**.** Las siguientes son las notas de estadística dos grupos de Corporación Universitaria Remington, de una muestra elegida al azar:

grupo 1

Xi (nota)	fai(n° estud)
1	5
2	10
3	8
4	3
5	10

grupo 2

Xi(nota)	fai(nº estud)
1	7
2	6
3	8
4	5
5	6

El entrenador de un equipo de atletismo está evaluando a tres estudiantes para poderlos incluir en su equipo. Se hicieron competir a estos 3 atletas en 5 carreras de 500 mts. Y se obtuvieron los siguientes resultados (en segundos):

Atleta A	61.3	62.1	61.7	62.9	63.2
Atleta B	62.8	63	62.5	61.9	60.7
Atleta C	61.9	63.7	62.9	61.9	61.5

5.4. Medidas de variabilidad o dispersión para datos cuantitativos

Para entender en que consiste la variabilidad, veamos el siguiente ejemplo:
A continuación se tiene el número de unidades producidas por hora durante un día por dos operarios:

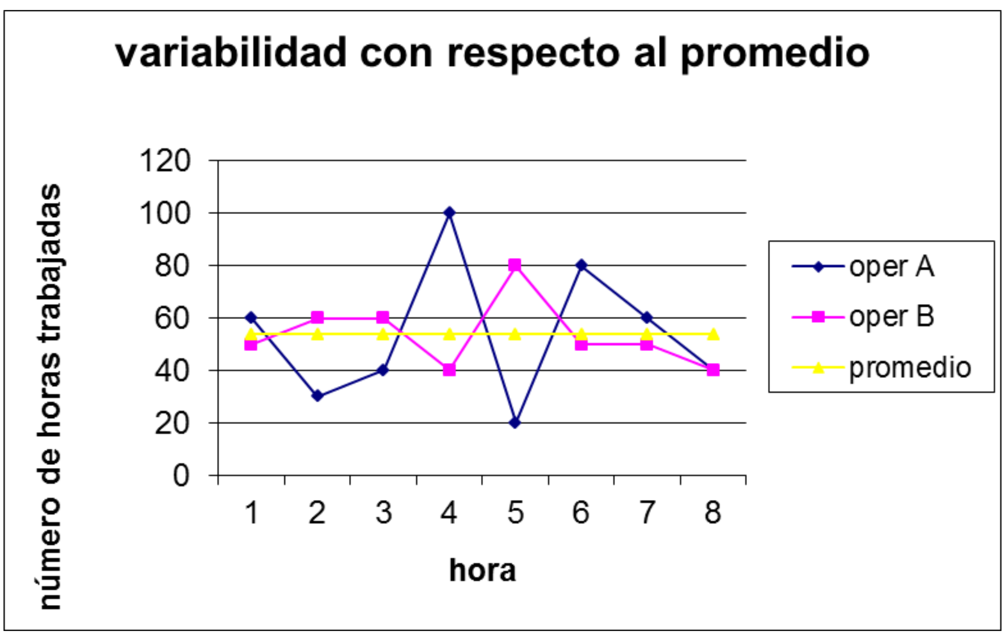
oper A	60	30	40	100	20	80	60	40
Oper B	50	60	60	40	80	50	50	40

Si vamos a buscar el promedio del número de unidades producidas por hora en el día de cada operario, sumamos las unidades y dividimos por las 8 horas así:

$\mu = \frac{60+30+40+100+20+80+60+40}{8} = \frac{430}{8} = 53.75$ para el operario A

$\mu = \frac{50+60+60+40+80+50+50+40}{8} \quad \frac{430}{8} = 53.75$ Para el operario B

Como podemos apreciar ambos operarios tienen el mismo promedio. Pero veamos algo adicional, vamos a trazar un diagrama de líneas para cada operario:



Como podemos ver, el operario A presenta mayor variación con respecto al promedio que el operario B. Las medidas que vamos a ver a continuación tienen dos aplicaciones:

1. Para analizar cómo varía un conjunto de datos con relación a su propio promedio.
2. Para comparar la variabilidad de dos o más conjuntos de datos entre sí.

Ahora veamos la primera de estas medidas:

1. **VARIANZA**

Si los datos no están organizados en una tabla de frecuencias, la varianza se calcula así:

Como parámetro, es decir, si los datos se toman de una población:

$$\sigma^2 = \frac{(x_1-\mu)^2 + (x_2-\mu 0)^2 + (x_3-\mu)^2 + \ldots + (x_n-\mu)^2}{N} = \frac{\sum(x_i-\mu)^2}{N}$$

Como estadística, es decir, si los datos se toman de una muestra:

$$S^2 = \frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + (x_3-\bar{x})^2 + \ldots + (x_n-\bar{x})^2}{n-1} = \frac{\Sigma(x_i-\bar{x})^2}{n-1}$$

Veamos el ejemplo que tenemos sobre los operarios:

OPERARIO A			OPERARIO B	
Xi	**$(Xi-\mu)^2$**		**Xi**	**$(Xi-\mu)^2$**
60	39,0625		50	14,0625
30	564,0625		60	39,0625
40	189,0625		60	39,0625
100	2139,0625		40	189,0625
20	1139,0625		80	689,0625
80	689,0625		50	14,0625
60	39,0625		50	14,0625
40	189,0625		40	189,0625
$\sum$=430	$\sum$=4987,5		$\sum$=430	$\sum$=1187,5

$\mu_A = \frac{\sum Xi}{N} = \frac{430}{8} = 53.75$ $\mu B = \frac{\sum Xi}{N} = \frac{430}{8} = 53.75$

$\sigma^2{}_A = \frac{\Sigma(Xi-\mu)^2}{N} = \frac{4987.5}{8} = 623.438$ $\sigma^2{}_B = \frac{\Sigma(Xi-\mu)^2}{N} = \frac{1187.5}{8} = 148.438$

Pero si los datos están organizados en una tabla de frecuencias, la varianza se calcula

Como parámetro, es decir, si los datos se toman de una población:

$$\sigma^2 = \frac{(x_1-\mu)^2.fa_1 + (x_2-\mu)^2.fa_2 + (x_3-\mu)^2.fa_3 + \ldots + (x_n-\mu)^2.fa_n}{N} = \frac{\Sigma(x_i-\mu)^2.fai}{N}$$

Como estadística, es decir, si los datos se toman de una muestra:

$$S^2 = \frac{(x_1-\bar{x})^2.fa_1 + (x_2-\bar{x})^2.fa_2 + (x_3-\bar{x})^2.fa_3 + \ldots + (x_n-\bar{x})^2.fa_n}{n-1} = \frac{\Sigma(x_i-\bar{x})^2.fai}{n-1}$$

OPERARIO A

Xi	**fai**	**Xi*fai**	**$(Xi - \mu)^2$*fai**
20	1	20	1139,0625
30	1	30	564,0625
40	2	80	378,125
60	2	120	78,125
80	1	80	689,0625
100	1	100	2139,0625
	N = 8	∑= 430	∑= 4987.5

$\mu = \frac{\sum Xi*fai}{N} = \frac{430}{8} = 53.75$

$$\sigma^2 = \frac{\Sigma(Xi-\mu)^2 * fai}{N} = \frac{4987.5}{8} = 623.438$$

OPERARIO B

Xi	fai	Xi*fai	(Xi - μ)^2*fai
40	2	80	378,125
50	3	150	42,1875
60	2	120	78,125
80	1	80	689,0625
	N = 8	∑= 430	∑= 1187.5

$$\mu = \frac{\sum Xi * fai}{N} = \frac{430}{8} = 53.75$$

$$\sigma^2 = \frac{\Sigma(Xi-\mu)^2 * fai}{N} = \frac{1187.5}{8} = 148.438$$

Pasos para calcular la varianza:

1. Se busca el promedio: μ ó $\overline{x}$
2. Se calcula la desviación de cada dato con respecto al promedio y este resultado se eleva al cuadrado: (xi-μ)2 ó (xi- $\overline{x}$)2
3. Cada uno de estos se multiplican por la frecuencia absoluta, en caso de que exista: $(x_i-\mu)^2.fa_i$ ó $(x_i-\overline{x})^2.fa_i$
4. Se suman los anteriores resultados.
5. Esta sumatoria se divide por el número de datos. N o n-1, según el caso.

Como podemos observar, la varianza da en unidades al cuadrado lo cual resulta ilógico y no se puede interpretar, por esta razón se creó la siguiente medida de variación:

2. DESVIACIÓN ESTÁNDAR

También conocida como desviación típica, "es la raíz cuadrada de la varianza. Expresa la dispersión de la distribución y se expresa en las mismas unidades de medida de la variable. La desviación típica es la medida de dispersión más utilizada en estadística".
(http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp#introduccion)

Como parámetro, es decir, si los datos se toman de una población:

$$\sigma = \sqrt{\frac{(x_1-\mu)^2 . fa_1 + (x_2-\mu)^2 . fa_2 + (x_3-\mu)^2 . fa_3 + ... + (x_n-\mu)^2 . fa_n}{N}}$$

Como estadística, es decir, si los datos se toman de una muestra:

$$S = \sqrt{\frac{(x_1-\bar{x})^2 . fa_1 + (x_2-\bar{x})^2 . fa_2 + (x_3-\bar{x}) . fa_3 + ... + (x_n-\bar{x})^2 . fa_n}{n-1}}$$

Para el ejemplo que tenemos de los operarios, la desviación estándar se calcula así:

Operario A $\sigma = \sqrt{623.438} = 24.969$

El número de unidades producidas por hora durante día del operario A tiene una desviación promedio de 24.964 unidades con relación a su promedio 53.75

Operario B $\sigma = \sqrt{148.438} = 12.184$

El número de unidades producidas por hora durante día del operario B tiene una desviación promedio de 12.184 unidades con relación a su promedio 53.75

La desviación estándar es útil siempre y cuando se estén comparando conjuntos de datos que tengan promedios similares y que sean las mismas unidades de investigación como en el ejemplo de los dos operarios, pero cuando los promedios y las unidades de investigación son diferentes, no basta con esta medida; por ejemplo si se está comparando el número de unidades producidas por hora de un operario y el sueldo de otro operario, no se podrían comparar las variaciones de estos.

Veamos un ejemplo:

Un operario C produce en promedio 40 unidades por hora, con una desviación estándar de 5.

Otro operario D produce en promedio 160 unidades por hora, con una desviación estándar de 15.
A simple vista parece ser que el operario D tiene 3 veces más variabilidad que el operario C; pero debe tenerse en cuenta que el operario D produce unidades en promedio 4 veces más que el operario C; así que para estos casos se tiene la siguiente medida:

1. **COEFICIENTE DE DISPERSIÓN O DE VARIACIÓN**

También conocida como variación relativa, puesto que muestra en qué porcentaje está variando un conjunto de datos; es decir, que no está expresado en las unidades de investigación.

C.V = $\frac{\sigma}{\overline{x}} * 100$ para una población C.V = $\frac{s}{\overline{x}} * 100$ para una muestra

En ejemplo anterior:

C.Vc = 5 * 100 =12.5% C.VD = 15 * 100 = 9.4%
160

Así podemos ver que el operario D presenta menor variabilidad.

Para el ejemplo que tenemos de los operarios A y B:

C.VA = $\frac{24.969}{53.75} * 100 = 46.454$ % C.V$_b$ = $\frac{12.184}{53.75} * 100 = 22.668$ %

Realicemos un ejemplo con un conjunto de datos agrupados:

A continuación se dan las edades de los empleados de una empresa:

LI	**LS**	**fai**	**Xi**	**Xi*fai**	**(Xi-μ)2* fai**
18	23	7	20,5	143,5	1411,08243
23	28	20	25,5	510	1692,06408
28	33	15	30,5	457,5	264,34806
33	38	7	35,5	248,5	4,502428
38	43	4	40,5	162	134,652816
43	48	25	45,5	1137,5	2917,0801

48	53	3	50,5	151,5	749,109612
		N =81		∑=2810.5	∑=7172.84

$$\mu = \frac{\sum Xi * fai}{N} = \frac{2810.5}{81} = 34.698$$ Este es el promedio de las edades de los empleados

$$\sigma^2 = \frac{(Xi-\mu)^2 * fai}{N} = \frac{7172.84}{81} = 88.554$$ No se interpreta

$$\sigma = \sqrt{\sigma^2} = \sqrt{88.554} = 9.41$$

$$C.V = \frac{9.41}{34.689} * 100 = 27.127$$ % Este es el porcentaje de variación

Las edades de los empleados de la empresa tienen una desviación promedio de 9.41 años, con relación a su promedio 34.689 años.

Ejercicios tema 4

El gerente de una empresa comercializadora tiene el record de las ventas de sus 3 vendedores durante los últimos 5 meses (en millones de pesos)

Mónica	88	68	90	101	89
Alex	77	89	90	87	78
Sandra	104	88	118	88	123

Calcule: Varianza, Desviación Estándar y coeficiente de Variación o Dispersión.

Actividad final unidad 3

Los siguientes son los ingresos semanales (en millones de pesos) de 2 centros de atención psicológica durante los últimos 2 años, de acuerdo con una muestra aleatoria:

Centro psicológico A

Li	Ls	fai
0	1	13
1	2	40
2	3	10
3	4	24
4	5	3

Centro psicológico B

Li	Ls	fai
0	1	33
1	2	10
2	3	10
3	4	40
4	5	7

Seleccione la respuesta correcta:

1. Los ingresos más frecuentes del Centro psicológico A están en el intervalo:

a) De 0 a 1 **b)** De 1 a 2 **c)** De 3 a 4

2. Los ingresos menos frecuentes del Centro psicológico A están en el intervalo:

a) De 2 a 3 **b)** De 4 a 5 **c)** De 2 a 3

3. Los ingresos más frecuentes del Centro psicológico B están en el intervalo:

a) De 4 a 5 **b)** De 1 a 2 **c)** De 3 a 4

4. Los ingresos menos frecuentes del Centro psicológico B están dos intervalos:

a) Falso b) verdadero

5. Los ingresos menores de los dos Centros psicológicos están en el intervalo:

a) De 0 a 1 **b)** De 1 a 2 **c)** De 3 a 4

6. Los ingresos mayores de los dos Centros psicológicos están en el intervalo:

a) De 4 a 5 **b)** De 1 a 2 **c)** De 3 a 4

7. Se puede afirmar que el Centro psicológico B tuvo el mayor promedio de ingresos:

a) Falso **b)** verdadero

Actividad

Realiza una investigación estadística, sobre datos cuantitativos agrupados, en tu medio; ya sea tu lugar de trabajo, tu ciudad o tu familia y realiza todo el proceso: tablas, gráficos, medidas y conclusiones o decisiones finales.

5.5. Glosario

Dato más frecuente: Es el dato que más se repite; es decir la moda. Se identifica como el que tiene la frecuencia absoluta más alta.

Frecuencias: Indica en forma numérica (absoluta) o en forma porcentual (relativa) las veces que se presenta cada dato.

Inferencia: Es la generalización que se obtiene, partiendo de una o varias muestras, sobre la población.

Dispersión: Indica cómo se dispersan o varían los datos en la distribución; existen varias medidas para analizar dicha dispersión; las más utilizadas son las que varían con relación al promedio.

Histograma: Es un gráfico de barras continuas y puede ser de frecuencias absolutas o frecuencias relativas**.**

Ojiva: Muestra gráficamente el comportamiento numérico o porcentual de los datos en la forma: "menor o igual que el dato"

Diagrama de barras: Es el que más se aplica en datos cuantitativos ordenados en fila.

Frecuencias acumuladas: Las frecuencias absolutas y las relativas, se acumulan por cada clase y se utilizan para hacer interpretaciones de los datos como: mayor o igual, menor, menor o igual.

Interpretación de datos: mayor, dato menor, dato más frecuente, dato menos frecuente. Consiste en el análisis de los datos con el fin de analizar el comportamiento de ellos y concluir.

5.6. Fuentes Bibliográficas

Anderson, D., Sweeney, D. & Williams, T. (1999). Estadística para Administración y Economía. (7ª edición). México: Internacional Thomson Editores.

Berenson, M. L. & LEVINE, D. M. (1996). Estadística básica en Administración. (6ª edición). México: Prentice-Hall.

Cáceres Hernández, J. (2009). Conceptos básicos de Estadística para ciencias sociales. Madrid: Delta Publicaciones.

Espejo, M. (2003). Estadística descriptiva y probabilística. Cádiz: Universidad de Cádiz.

Martínez Bencardino, C. (2004). Estadística y muestreo. (11ª edición). Bogotá: Ecoe ediciones.

Mendenhall, W. & Sincich, T. (1997). Probabilidad y Estadística para Ingeniería y administración. (4ª edición). México: Prentice-Hall.

Pérez López, J. (2007) Muestreo estadístico. Madrid: Prentice-Hall.

Ross, Sheldon, M. (2005). Introducción a la Estadística. Barcelona: Reverte.

Spiegel, M. R. (1995).Estadística. (2ª edición). Madrid: McGraw-Hill.

Berenson, M. L. & Levine, D. M. (1996). Estadística básica en Administración. (6ª edición). México: Prentice-Hall.

Anderson, D., Sweeney, D. & Williams, T. (1999). Estadística para Administración y Economía. (7ª edición). México: Internacional Thomson Editores.

Spiegel, M. R. (1995).Estadística. (2ª edición). Madrid: McGraw-Hill.

5.7. Fuentes Digitales o Electrónicas

Compas3 Comercio Electrónico. Introducción a la estadística descriptiva [Versión electrónica]. Madrid, España, 2000. Extraído el 10 de octubre de 2009 de: http://www.aulafacil.com/CursoEstadistica/CursoEstadistica.htm

María Da Silva Ramis. Definición y Aplicaciones de la estadística descriptiva [Versión electrónica]. Extraído el 27 de octubre de 2009 de:

http://www.monografias.com/trabajos10/esta/esta.shtml?monosearc

Pita Fernández, S. Estadística descriptiva de los datos Uso de la estadística y la epidemiología en atención primaria. En: Gil VF, Merino J, Orozco D, Quirce F.

Manual de metodología de trabajo en atención primaria. Universidad de Alicante. Madrid, Jarpyo Editores, S.A. 1997; 115-161. Actualizado 06/03/2001. Extraído el 27 de octubre de 2009 de:http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp#introduccion

Universidad de San Carlos. Estadística descriptiva: Conceptos básicos [Versión electrónica]. Guatemala, actualizado el 21de agosto de 2007. Extraído el 24 de octubre de 2009 de http://sitios.ingenieria-usac.edu.gt/estadistica/estadistica2/estadisticadescriptiva.html