



COLEGIO DE BACHILLERES

ESTADÍSTICA DESCRIPTIVA E INFERENCIAL I

FASCÍCULO 1. ANTECEDENTES HISTÓRICOS DE LA
ESTADÍSTICA Y SUS FUNCIONES

Autores: Hugo García Mancilla
Juan Matus Parra



Colaboradores

Asesoría Pedagógica
Irma Cruz Santillán

Revisión de Contenido
Armando Martínez Cruz

Diseño Editorial
Leonel Bello Cuevas
Javier Darío Cruz Ortiz

ÍNDICE

INTRODUCCIÓN	5
PROPÓSITO	7
CAPÍTULO 1 ANTECEDENTES HISTÓRICOS DE LA ESTADÍSTICA Y SUS FUNCIONES	9
1.1 ANTECEDENTES HISTÓRICOS DE LA ESTADÍSTICA Y SUS FUNCIONES	9
1.1.1 Bosquejo Histórico	9
1.1.2 Funciones de la Estadística	14
1.1.3 Aplicaciones	18
1.1.4 Conceptos preliminares	20
1.1.5 Clasificación de la Estadística	26
1.1.6 Estadística Descriptiva	27
1.1.7 Estadística Inferencial	28
1.2 CONCEPTOS BÁSICOS PARA EL ESTUDIO DE LA ESTADÍSTICA	30
1.2.1 Presentación de Datos	37
1.2.2 Distribución de Frecuencias	37
1.2.3 Distribución de Frecuencias Absolutas y Relativas	39
1.2.4 Distribución de Frecuencias Acumuladas	52

1.2.5	Gráficas	54
1.2.6	Histograma y Polígono de Frecuencias	55
1.2.7	Polígono de Frecuencia	59
1.2.8	Polígono de Frecuencia Acumulada	62
1.2.9	Gráficas circulares, Diagramas de Barras, Pictogramas y Gráfico de líneas.	62
RECAPITULACIÓN		74
ACTIVIDADES DE CONSOLIDACIÓN		75
AUTOEVALUACIÓN		78
BIBLIOGRAFÍA CONSULTADA		87

INTRODUCCIÓN

La estadística es una de las herramientas más ampliamente utilizadas en la investigación científica.

Su aplicación en instituciones gubernamentales y educativas, en los negocios y en la industria, en la banca y en otros quehaceres diarios hacen de la estadística una herramienta indispensable.

Sin embargo el término “Estadística” tiene varios significados para diferentes personas; para la gente común y corriente la estadística solamente significa números. En el periodo de la mañana se pueden encontrar la estadística más reciente sobre los delitos de la ciudad; de asesinatos, de robos de automóviles; de asaltos y demás delitos que hayan sido denunciados en determinado periodo de tiempo; de los nacimientos y muertes que han ocurrido, o en relación con el deporte, el número de partidos ganados y perdidos por equipos integrantes de la liga de ese deporte.

Para otras personas es un método para obtener, presentar y escribir grandes cantidades de datos, y para otras es un método para tomar decisiones en situaciones difíciles.

El objetivo básico de este fascículo es aclarar los significados de Estadística, definir sus conceptos básicos utilizados con frecuencia y analizar los usos y abusos de los métodos estadísticos.

Aunque los significados sean diferentes, todos ellos forman parte del concepto total de “Estadística”. La palabra tiene su sentido más amplio para aquellas personas cuyo trabajo requiere un conocimiento de los aspectos más técnicos de la estadística. Para estas personas, la palabra tiene relación con aquellos conceptos y técnicas que se utilizan en la recopilación, organización, resumen, análisis, interpretación y comunicación de información numérica.

Estos conceptos y técnicas juegan un papel importante en las actividades que realizan los profesionales de todas las ciencias.

PROPÓSITO

La estadística es una rama de las matemáticas aplicadas que surgió por la necesidad concreta que el hombre tiene de conocer la resolución de problemas relacionados con la recolección, procesamiento, análisis e interpretación de datos numéricos cuyo conocimiento le permitirá tomar decisiones acertadas.

Para el conocimiento de la realidad concreta que al hombre le interesa, considera tres etapas fundamentales que son:

- Planear la búsqueda y la obtención de la información.
- Sistematizar y organizar la información de tal forma que se pueda describir y analizar con facilidad.
- Efectuar inferencias sobre la realidad a partir de la información obtenida, haciendo estimaciones o verificando hipótesis.

La interpretación de la información permite obtener conclusiones que enriquecen nuestro conocimiento de la realidad y nuestra capacidad para transformarla.

El propósito de este fascículo es el de proporcionarte los conocimientos necesarios para llevar a la práctica las etapas que te permitirán la resolución de cualquier problema estadístico.

CAPÍTULO 1

ANTECEDENTES HISTÓRICOS DE LA ESTADÍSTICA Y SUS FUNCIONES

1.1 ANTECEDENTES HISTÓRICOS DE LA ESTADÍSTICA Y SUS FUNCIONES

Al igual que ha ocurrido con otras muchas disciplinas, a lo largo del tiempo se ha pensado que la estadística es un procedimiento extraordinariamente complicado. Cuando leemos artículos en los que aparecen resultados estadísticos nos queda la impresión de que lo dicho en ellos es una verdad absoluta e incontrovertible que está apoyada por todo un aparato matemático. Esto no es forzosamente cierto, se puede adquirir con relativa facilidad un conocimiento básico de la estadística.

1.1 Bosquejo Histórico

Iniciaremos el estudio de la estadística con algunos antecedentes históricos, que nos mostrarán sus aplicaciones, por una parte y por otra, su naturalidad en situaciones de la vida real.

La estadística fue fundada por el londinense John Graunt, “un mercader de mercería”, en un pequeño libro “Natural and political Observations made upon the Bells of Mortality”. Este libro fue el primer intento para interpretar fenómenos biológicos de masa y de la conducta social: a partir de datos numéricos escribir las cifras brutas de nacimientos y defunciones en Londres, de 1604 a 1661. El opúsculo de Graunt apareció en 1662. Treinta años más tarde, la Royal Society publicó en su “Philosophical Transactions” un artículo sobre tasas de mortalidad escrito por el eminente astrónomo Edmund Halley. Ambas publicaciones constituyen la base de todo trabajo posterior sobre esperanza de vida, indispensable para la solvencia de las compañías de seguros de vida.

John Graunt nació en 1620 en Berchin Lane, Londres, bajo el signo de las siete estrellas, donde su padre tenía una tienda y el hogar. Aprendió pronto el oficio de vendedor de mercería y prosperó en el negocio. El éxito le dio la posibilidad de dedicarse a ocupaciones más amplias que las de la venta de artículos de mercería. Aubrey lo describe como “una persona muy ingeniosa y estudiosa... se levantaba muy temprano para sus estudios antes de abrir la tienda”. Se hizo amigo de Sir William Petty, más tarde autor de un conocidísimo libro sobre la nueva ciencia de la aritmética política, y probablemente discutió con él las ideas expresadas en sus “Observations”.

Las tablas de mortalidad, que atrajeron la atención de Graunt, eran publicadas semanalmente por la compañía de Sacristanes parroquiales y contenían el número de muertes acaecidas en cada parroquia, sus causas y también un “Recuento de todos los entierros y bautizos habidos en la semana” en las cuales anotaban el número de nacimientos de acuerdo a los que acudían al bautismo y lo mismo sucedía cuando presentaban sus defunciones (en las parroquias se llevaba el control).

Un ejemplo de las observaciones hechas por Graunt en 1632 fueron las siguientes:

Varones	4,994
Bautizados Hembras	4,590
T o t a l	9,584

Varones	4,932
Enterrados Hembras	4,603
T o t a l	9,535

Con estos datos deducía las siguientes observaciones:

- a) Hay más varones que hembras
- b) Pocos murieron de hambre
- c) Hay pocos asesinatos
- d) Los lunáticos son pocos

Las “Observations” impresionaron tan favorablemente a Carlos II, que este propuso especialmente a Graunt como socio fundador de la recientemente constituida Royal Society. Para prevenir cualquier posible objeción al hecho de que Graunt era tendero, “su majestad dio este encargo particular a su Sociedad, de que si encontraban algún comerciante más de su estilo, lo admitiesen sin más ceremonia”. Graunt fue elegido socio fundador de la Royal Society en 1662.

El mérito de las “Observations” fue inmediatamente reconocido, y fomentó el estudio de las estadísticas de vida en el continente. El libro alcanzó varias ediciones. La quinta, publicada tras la muerte de Graunt fue ampliada por Petty. Los historiadores han discutido largo tiempo la contribución de Petty al trabajo original. Aubrey que era malicioso, sólo dice que Graunt fue “inspirado” por Petty, pero implica mucho más. Parece indudable que el libro es una obra conjunta.

Desde luego, Graunt escribió la mayor parte, incluidas las aportaciones científicas más valiosas. Petty añadió lo que Thomas Browne llamaría “Elegancia”, y así aumentó la popularidad del libro. Sir William Petty era un hombre presuntuoso y algo engreído, incapaz de decidir si patrocinar a Graunt o acreditar su trabajo. No hay pruebas de que alguna vez hubiese entendido la importancia y originalidad de lo que había hecho su amigo.

Graunt fue miembro del consejo común de la ciudad y desempeñó otros cargos, pero al convertirse al catolicismo dejó el comercio y cualquier otra obra pública. Graunt tenía cabeza y talento para el trabajo, y era jocosos y fecundo en su conversación.

Graunt murió de ictericia la víspera de Pascua en 1674 y fue enterrado en la iglesia de St. Dunston.

John Arbuthnot ¹



En los trabajos de Graunt y Halley se basó John Arbuthnot en 1690 para probar la existencia de Dios. Su argumento dice:

No es posible la suposición de que el sexo está distribuido entre la descendencia humana en una forma puramente casual; debe intervenir una providencia divina que controla las proporciones de los sexos.

La demostración de Arbuthnot es el primer ejemplo conocido de inferencia estadística.

Anchenwall un economista, acuñó en 1760 la palabra estadística, que deriva del término italiano statista. La raíz de la palabra procede del latín status que significa estado o situación.

¹ www_history.mcs.st-andrews.ac.uk/history/pictDisplay/Arbuthnot.html

La Ley de los Grandes Números

En el famoso libro de Jacob Bernoulli, *Aos Conjectandi*, aparece un teorema de importancia cardinal para la Teoría de Probabilidades, comúnmente llamado Teorema de Bernoulli, y también conocido como Ley de los grandes números, nombre que le fue dado por el matemático francés, Simeon Poisson (1781-1840). Este teorema fue el primer intento para deducir medidas estadísticas a partir de probabilidades individuales. El tiempo empleado para escribir este libro no fue perdido, si consideramos la importancia central del resultado. Matemáticos, científicos y filósofos han dedicado más de veinte años examinando y discutiendo el significado exacto del Teorema y su alcance en aplicaciones estadísticas.

El teorema es más sencillo de exponer. De hecho, cuando se ve por primera vez, uno se pregunta cómo Bernoulli pudo preocuparse durante veinte años y cómo ha promovido tantas controversias posteriormente. El hecho es, que es un conjunto de sutilezas y artificios; cuando más lo piensa uno, más complicado lo ve. Bernoulli tuvo un trabajo loco montando el engranaje, lo cual lo distrajo de prever los embrollos lógicos y filosóficos que planteaba.

“Si la probabilidad de un suceso es p , y si se hace un número infinito de pruebas, la producción de aciertos es, sin duda p ”. Aquí, tienen una simple exposición del Teorema de Bernoulli: si la probabilidad de que ocurra un hecho en una prueba única es p , y si se hacen varias pruebas, inmediatamente y en las mismas condiciones, la proporción más probable de que ocurran los hechos en el número total de pruebas es también p ; aún más, la probabilidad que la porción en cuestión difiere de p en menos que una cantidad dada, por pequeña que sea, aumenta al mismo tiempo que aumenta el número de pruebas.

Tirando al aire su discreción matemática “un estudioso del sujeto llega a esta definición correcta”.

Otra definición más válida: “En un conjunto bastante amplio de “a” elementos es casi seguro que la frecuencia relativa de “b” elementos se aproximará a la probabilidad de un elemento “a” estando “b” dentro de cualquier grado de aproximación deseada”. Aquí la frase “casi seguro” ha de entenderse como un medio conveniente para decir que hay una probabilidad tan cercana como queramos a 1.

Como una demostración de la importancia de la Ley de los grandes número en asuntos prácticos es suficiente mencionar los Seguros.

Supongamos que la probabilidad de que un hombre de cierta edad y constitución muera en el transcurso de un año es $1/10$. si tal individuo decide asegurarse, ésta es la fracción que ha de tener en cuenta y usar cuando tome su decisión. Pero la compañía de seguros que se ofrece a cubrir el riesgo de su muerte en este período tiene en consideración otra probabilidad que se deriva de esta probabilidad. Si hay un gran número de personas de las mismas características, que aseguran sus vidas en esa compañía, hay una probabilidad muy elevada de que la compañía no tenga que pagar a más de, aproximadamente, un décimo de las pólizas. Si, por consiguiente, la compañía carga en cada caso una prima de más de un décimo del total de la póliza, es muy probable que tendrá bastante superávit después de pagar todos los derechos, para cubrir los gastos administrativos y distribuir un dividendo a sus accionistas.

Mientras mayor sea el número de personas que se asegura en la compañía, mayor es la probabilidad de que las finanzas de la compañía sean sanas siempre que las primas estén calculadas como acabamos de decir. Esta es la consideración fundamental que distingue el negocio de una compañía de seguros de una apuesta.

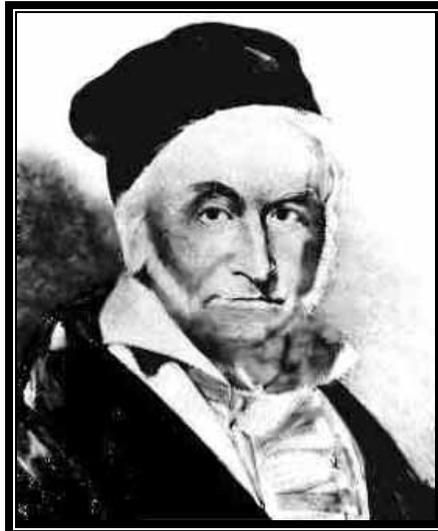
Girolamo Cardano ²



Cardano nació en Pravia en 1501 y murió en 1576. Su vida es una serie de actos incoherentes que pertenecen tanto a la historia de la Matemática como a la de la Astrología y a la de la Patología. Realizó sus primeros estudios en su ciudad natal y luego en la Universidad de Padua, donde alcanzó la Licenciatura en Medicina que ejerció en Sacco y en Milán durante el período de 1524 a 1556. Durante estos años estudió Matemáticas y publicó sus principales obras. Entre estas destaca el *Ars Magna*, en la cual se presentan raíces negativas de una ecuación, algunos cálculos con números imaginarios y la fórmula de la ecuación cúbica que ha pasado a la historia con el calificativo de Cardámica, aunque ya se sabe que es de Fortaglia, con quien tuvo una de las polémicas más agrias en la historia de las Matemáticas. Se le atribuye la primera discusión sobre “Probabilidad” en su manual para jugadores “*Siber De Ludo Aleae*” (Manual para tirar dados).

² www_history.mcs.st-andrews.ac.uk/history/pictDisplay/Cardan.html

Karl Fredrich Gauss – (1777-1855) –³



Junto con Arquímedes y Newton, Gauss es uno de los tres grandes de la Matemática. Ellos aportaron conceptos muy útiles en sus distintas ramas tanto en su forma pura como aplicada.

La precocidad de Gauss fue evidente antes de los tres años de edad. Cuando su padre hacía la nómina para pago de los trabajadores, sin darse cuenta que su hijo seguía sus acciones, al terminar, el niño exclamó “Padre el cálculo está equivocado”. Al comprobarlo notó que el resultado que le dijo el niño era correcto.

Gauss se hizo notable, ya que a los doce años criticó los fundamentos de la Geometría Euclídeana, a los trece le interesaba la posibilidad de la Geometría No-Euclídeana, a los quince entendió el concepto de convergencia de líneas y probó el binomio de Newton, a los dieciocho inventó el método de los mínimos cuadrados, a los diecinueve, el 30 de marzo de 1796, descubrió la construcción del polígono de 17 lados sólo con regla y compás.

La ley de Gauss de la distribución normal de errores y su curva en forma de campana usada por maestros, estadistas, comerciantes, etcétera, se denomina también curva normal de frecuencias y encuentra sus raíces en la Teoría Matemática de los juegos de azar.

Su lema fue: *Pauca. Sed natura*, que significa: “Poco, pero maduro”.

³ www_history.mcs.st-andrews.ac.uk/history/pictDisplay/Gauss.html

Johann Von Neumann – (1909-1957) –⁴



Epistemólogo austriaco contemporáneo. VON Neumann llevó a cabo la primera demostración del Teorema Minimax, base fundamental de la Teoría de juegos, que fue propuesto primeramente por Emile Borel en 1921. También fue pionero de la Teoría de Computadoras, habiendo diseñado y construido el llamado MANIAC (analizador matemático, integrador numérico y computador) en el Instituto para estudios avanzados de Pinceton, en 1952.

Sus ideas fundamentales sobre la axiomatización de las matemáticas las ha expuesto en varias memorias especialmente en *Eine Axiomatisierung der Mengenlehre*, *Crelle*, 1925 y *Axiomatisierung der Mengenlehre Math, Zaitrehj* 1928.

La situación actual de la Estadística se debe al esfuerzo de grandes matemáticos y científicos. Entre los más famosos se puede mencionar a Laplace, Fermat, Jacques, Bernoulli y Gauss, quienes intervinieron en el primer y más importante estudio de la probabilidad en los siglos XVIII y XIX.

El matemático belga Quetelet, los estadísticos escandinavos Charlier y Gram, los ingleses Pearson, Fisher, Galton, también asociaron sus nombres al progreso de esta nueva disciplina, a la que dotaron de bases matemáticas sólidas.

Como se observa, la evolución de la Estadística estuvo conformada por una serie de necesidades que condujeron al hombre a su creación.

⁴ www_history.mcs.st-andrews.ac.uk/history/pictDisplay/Von_Neumannn.html

1.1.2 Funciones de la Estadística

¿Para qué estudiamos estadística?

Los conceptos y temas de la estadística se utilizan en la actualidad en un gran número de ocupaciones. Las técnicas estadísticas constituyen una parte integral de las actividades de investigación en distintas áreas del saber humano.

La persona que comprenda los conceptos estadísticos y su metodología obtendrá mejor provecho de ellos. Su trabajo tal vez no necesite conocer la estadística sino aquello que lo faculte para saber cuándo se requieren los servicios de un experto y para poder comunicarse eficazmente con él.

El profesional, que entiende de estadística puede leer con inteligencia la literatura que sobre su campo de acción va apareciendo día con día.

Con frecuencia escuchamos en los medios de difusión comentarios como los siguientes:

- Se ha demostrado estadísticamente que el mayor porcentaje de las ventas de automóviles se registran en el primer trimestre del año.
- La explotación de petróleo crudo en el último trimestre del año de 1993 ascendió a 285 millones de barriles, cuyo producto fue de 3698 millones de dólares.
- Estadísticamente se ha demostrado que el huevo produce el colesterol en las personas que consumen mucho este producto.
- Se ha comprobado estadísticamente, que la pasta dental de mayor aceptación por el público es la que produce la fábrica Colgate-Palmolive.

Todas estas expresiones nos indican que la Estadística es una herramienta que ayuda a conocer la realidad. Sin embargo, también puede servir para distorsionar la verdad si no se tiene cuidado al usar los métodos estadísticos adecuadamente y si la interpretación de los resultados lo hacen incorrectamente.

Es famoso la frase que en cierta ocasión el ministro inglés Benjamin Disraeli dijo: Hay tres clases de mentiras que son: "Las simples, las malvadas y las de la Estadística".

Esta acusación hecha hace muchos años, ha llegado a convertirse en una descripción adecuada de algunos engaños que se pueden realizar mediante la Estadística.

Sucede algunas veces que en la presentación de resultados de alguna investigación poco seria, se utiliza la Estadística para dar la impresión de que todo ha sido realizado

en forma científica y al comprobarlo, resulta que la información fue alterada con la finalidad de obtener un resultado favorable pero falso.

Se debe tener cuidado en los reportes estadísticos en no caer en una falacia; para evitarlo se debe tener una actitud crítica hacia cualquier escrito que leamos. Así mismo hacer un análisis cuidadoso que nos indique si el contenido y la interpretación son correctas o si se ha hecho un mal uso de la Estadística.

Para un análisis crítico de la información contenido en un reporte, debemos contestarnos las siguientes preguntas:

1) ¿Cómo es la muestra?

La muestra debe ser representativa de la población en estudio.

2) ¿Qué se está midiendo y cómo?

Es importante saber de qué manera se obtienen los datos y si las características de interés pueden ser medidas.

Ejemplo:

Se desea saber qué tiempo en horas un alumno dedica al estudio, para ello el encuestador le hace la siguiente pregunta:

Tú estudias dos horas diarias, ¿verdad?

Esta forma de obtener la información es incorrecta toda vez que la respuesta es guiada por el encuestador y la mayoría van a ser afirmativas.

3) ¿Qué se está probando y cómo?

Si el reporte incluye la prueba de alguna hipótesis, entonces hay que buscar el planteamiento explícito de ésta; haciendo uso de los métodos estadísticos apropiados y en especial los inferenciales.

4) ¿Se está hablando siempre de lo mismo?

Con frecuencia se usan datos reales para obtener conclusiones falsas.

En algunos estudios las trampas de este tipo pueden estar en proceso de medición, veamos el siguiente.

Ejemplo:

En un poblado del Edo. De México se levantó un censo de todos los habitantes que no sabían leer para inscribirlos en los cursos de INEA. El encuestador reportó 3000 analfabetos que fueron los que detectó, pero él no sabía que mucha de esta gente no querían aprender a leer y por lo tanto no se presentaron para su reporte.

En el mismo poblado se anunció por la radio que toda persona que no supiera leer, el programa de “solidaridad” le otorgaría \$ 50,000 pesos mensuales de ayuda por lo tanto debían anotarse en la presidencia municipal las personas que cubrieran el requisito.

A este anuncio acudieron 7,000 personas que no sabían leer. Esto nos muestra que la información no siempre va a ser verdadera y en casos como estos, depende del interés de la gente.

5) ¿Tiene sentido la información?

Esta pregunta nos invita a analizar la información reportada y nosotros debemos ver si es congruente con el sentido común.

Ejemplo:

Un estudio reporta que la cosecha de frijol en el estado de Hidalgo está en relación directa con la precipitación pluvial, es decir, si llueve más, hay mayor cosecha. Esto es cierto en determinados límites toda vez que con demasiada lluvia se arruina la cosecha.

6) ¿Qué información falta?

Existe información que no es estrictamente estadística pero es necesaria para poder captar la que nos interesa por estar relacionada con ella.

Ejemplo:

En una librería se reportó que en tan sólo un mes se duplicaron las ventas, sin embargo, no se especificó que se refiere al mes de septiembre que fue el inicio de clases y además aumentó el costo de los libros debido a la demanda.

7) ¿Quién lo dice?

Muchas de las estadísticas que dan los medios de difusión se basan en la publicidad.

Ejemplo:

Un anuncio de la televisión dice: “K2 su fábrica amiga en tan sólo un mes vendió 50,000 colchones; pero aún nos quedan 1,000 que se rematan a un precio muy bajo, venga y llévase el suyo. ¿Será cierta esta información, o solamente es un comercial sensacionalista?

8) ¿Es correcta la representación gráfica que nos muestra la información?

Las representaciones gráficas también pueden ser engañosas toda vez que algunas veces se toman dos escalas diferentes para graficar el fenómeno, una para el eje horizontal y otra para el eje vertical, y si esto no se analiza con cuidado, se puede caer en una inferencia falsa.

De todo lo anterior podemos concluir que algunos reportes, ya sean gráficos o escritos que incluyen un informe estadístico, puede mentir; pero esto no quiere decir que siempre sea así.

Generalmente la Estadística se usa correctamente para poder obtener inferencias verdaderas.

Para no dejarte engañar es necesario que conozcas y manejes los aspectos de la Estadística y analizar con actitud crítica toda clase de información estadística.

ACTIVIDAD DE REGULACION

Recorta de periódicos o revistas los artículos referentes a información estadística. Lo podrás reconocer por la presencia de gráficos de números, como promedios, porcentajes, etcétera.

Analiza los recortes obtenidos y escribe en tu cuaderno las siguientes observaciones:

- a) ¿Qué variables se estudian de la población?
- b) ¿Qué características se reportan?
- c) ¿Cómo crees que se obtuvo la información?
- d) ¿Qué inferencia o conclusión se desprende del trabajo estadístico reportado?

Para cualquier duda consulta con tu profesor o asesor quien te indicará si tus respuestas son correctas.

1.1.3 Aplicaciones

¿En qué áreas se aplica la Metodología Estadística?

La metodología estadística se emplea en muchos campos. Se ha visto que la estadística es una disciplina que ayuda a diseñar el esquema de búsqueda y registro de información para describirla y analizarla con facilidad y mediante estimaciones, obtener conclusiones que enriquecen el conocimiento de la realidad.

La estadística día a día gana terreno en su aplicación en toda actividad humana por simple que ésta sea.

La estadística se aplica en los programas de Gobierno, Ingeniería, Agronomía, Economía, Medicina, Biología, Psicología, Pedagogía, Sociología, Física, etcétera; no hay alguna ciencia que no la use o profesión que no la aplique.

Algunos ejemplos del uso de la estadística son:

- 1) En las agencias gubernamentales, tanto federales como estatales utilizan la estadística para realizar planes y programas para el futuro.
- 2) En el campo de la ingeniería se aplica en muchas de sus actividades tales como:
 - a) La planeación de la producción.
 - b) El control de calidad.
 - c) Las ventas.
 - d) El almacén, etcétera.
- 3) En la Sociología se aplica para comparar el comportamiento de grupos socioeconómicos y culturales y en el estudio de su comportamiento.
- 4) En el campo económico su uso es fundamental para informar el desarrollo económico de una empresa o de un país que da a conocer los índices económicos relativos a la producción, a la mano de obra, índices de precios para el consumidor, las fluctuaciones del mercado bursátil, las tasas de interés, el índice de inflación, el costo de la vida, etcétera.

Todos estos aspectos que se estudian, se reportan e informan, no solamente describen el estado actual de la economía sino que trazan y predicen el camino de las futuras tendencias. Así mismo sirve a los encargados de las agencias, para tomar decisiones acertadas en sus operaciones.

- 5) En el campo demográfico la Estadística se aplica en los registros de los hechos de la vida diaria, tales como:
 - Nacimientos.
 - Defunciones.
 - Matrimonios.
 - Divorcios.
 - Adopciones.
 - Etcétera.

En materia de población los datos aportan una buena ayuda para fijar la política de estímulos al control de la natalidad, dirigir la inmigración o emigración, establecer los planes de lucha contra las enfermedades epidémicas o plagas que azotan los campos, etcétera.

- 6) En el campo educativo la Estadística contribuye al conocimiento de las condiciones fisiológicas, psicológicas y sociales de los alumnos y de los profesores. Al perfeccionamiento de los métodos de enseñanza y de evaluación.
- 7) Industria. La mayor parte de los industriales la utilizan para el control de calidad.
- 8) Agricultura. Se emplea en actividades como experimentos sobre la reproducción de plantas y animales entre otras cosas. También se usa la Estadística para determinar los efectos de clases de semillas, insecticidas y fertilizantes en el campo.
- 9) Biología. Se emplean métodos estadísticos para estudiar las reacciones de las plantas y los animales ante diferentes períodos ambientales y para investigar la herencia. Las leyes de Mendel sobre la herencia en donde los factores hereditarios se atribuyen a unidades llamadas genes y al estudio sistemático de los cruzamientos entre individuos portadores de genes diferentes, lo que ha permitido precisar de qué manera los genes se separan o se reúnen en las generaciones sucesivas. La verificación de las hipótesis formuladas por Mendel y sus continuadores necesitó el empleo de la Estadística, la cual en este caso ha lanzado las primeras luces sobre el mecanismo de la herencia.
- 10) Medicina. Los resultados que se obtienen sobre efectividad de fármacos se analizan por medio de métodos estadísticos. Los médicos investigadores se ayudan del análisis estadístico para evaluar la efectividad de tratamientos aplicados. La Estadística también se aplica en el establecimiento y evaluación de los procedimientos de medida o clasificación de individuos con el propósito de establecer la especificidad y sensibilidad a las enfermedades.
- 11) Salud. Los técnicos de la salud la utilizan para planear la localización y el tamaño de los hospitales y de otras dependencias de salud. También se aplica en la investigación sobre las características de los habitantes de una localidad, sobre el diagnóstico y la posible fuente de un caso de enfermedad transmisible; sobre la proporción de personas enfermas en un momento determinado, de ciertos padecimientos de una localidad, sobre la proporción de enfermos de influenza en dos grupos, uno vacunado contra el padecimiento y el otro no. También se aplica en cualquier otro tipo de investigación similar a éste.
- 12) Psicología. Los psicólogos se valen de los conceptos y técnicas de la estadística para medir y comparar la conducta, las actitudes, la inteligencia y las aptitudes del hombre.
- 13) Negocios. Los hombres de negocios pueden predecir los volúmenes de venta, medir las reacciones de los consumidores ante los nuevos productos, etcétera.

14) En la Física se utiliza la Estadística para obtener datos y probar hipótesis.

1.1.4 Conceptos Preliminares

¿Qué es la Estadística?

La mayor parte de las palabras tienen varios significados. La palabra “Estadística” no es una excepción. En el lenguaje común, la palabra se emplea para denotar un conjunto de calificaciones o de números, por ejemplo: una persona puede preguntar “¿has visto las últimas estadísticas acerca del desempleo?”. El comentarista de deportes puede decir “estas son las estadísticas para la primera mitad del partido”, o “¿has leído las estadísticas de accidentes en carreteras durante las vacaciones?”. El término estadística empleado así significa más que datos numéricos y no debe confundirse con la misma palabra empleada en este fascículo.

El término “estadística” también se emplea para designar un área de estudio: una disciplina. Como área de estudio, la Estadística proporciona los métodos que ayudan a resolver los problemas correspondientes, para ello es necesario conocer los conceptos básicos que conforman los contenidos de esta materia tales como:

- a) Población.
- b) Muestra.
- c) Variable de Estudio.
- d) Datos u observaciones.
- e) Parámetros.
- f) Estadígrafos o Estadísticos.

Veamos el siguiente ejemplo:

En el plantel 11 del Colegio de Bachilleres se aplicó una encuesta a 50 profesores de las diferentes áreas acerca de la adquisición anual de libros de la materia que imparten; para mantenerse actualizados.

Los resultados se muestran en la siguiente tabla:

No. De Libros	No. De Profesores
0	4
1	5
2	6
3	10
4	8
5	7
6	6
7	4
Total	50

La población en este caso son todos los profesores del Plantel 11. Cada uno de los profesores seleccionados y que se les aplicó la encuesta es un elemento, un dato o una observación y todos ellos forman una muestra representativa de la población.

Veamos otro ejemplo:

Iván Jerónimo es un alumno de la UNAM y ha participado en los concursos deportivos que la Institución organiza cada año. Iván es un corredor de los cien metros planos y cada vez que corre se impone un nuevo récord.

Los tiempos que ha establecido se muestran en la siguiente tabla:

No. de Participaciones	Tiempo en segundos
1	11.2
2	10.8
3	10.7
4	10.6
5	10.5
6	10.4
7	10.3

Analiza el ejemplo de los maestros y en base a estos conceptos, contesta en la línea las siguientes preguntas de este ejemplo:

- ¿Cuántas observaciones se tienen? _____
- ¿Cuántos datos son del problema? _____
- ¿Cuántos elementos tiene la muestra? _____
- ¿Cuál sería la población? _____

De los ejemplos anteriores podemos ver que al realizar un experimento para conseguir información acerca de un problema, se tiene un conjunto de resultados en que cada uno corresponde a una observación.

Ejemplo:

Se desea conocer la estatura de cada alumno del Plantel 2 del C.B.

El número de estudiantes que tiene el Plantel 2 es de 8,000 alumnos.

De este ejemplo contesta lo siguiente:

- ¿Cuántos elementos tiene la población? _____
- ¿Cuál es el número máximo de observaciones? _____

- ¿Cuál es el número mínimo de observaciones? _____

- Para no tener que medir a todos ¿qué propondrías? _____

De los ejemplos anteriores habrás notado que los datos se obtienen mediante la observación del fenómeno que interesa estudiar. También se obtienen de los documentos donde se reportan los hechos.

Veamos el siguiente ejemplo:

Se desea conocer el número de pacientes a quienes se les tomó placas de rayos X en el Hospital "Primer de Octubre", en el primer trimestre del año de 1993.

Para recabar esta información hay varias formas que son:

- Preguntarle al médico de guardia.
- Recurrir al expediente de los enfermos.
- Recurrir al informe de los médicos.
- Recurrir al informe del operador de la máquina de rayos X.

Contesta en la línea a ¿cuál de los anteriores recurrirías? _____

¿Por qué? _____

La información la podemos obtener en las diferentes instituciones que diariamente reportan su información tales como:

- Los bancos registran datos importantes.
- Las Iglesias registran nacimientos, defunciones y otros.
- Las Escuelas también registran datos importantes que en cualquier momento puede servir para realizar la investigación deseada.

Actualmente el Instituto Nacional de Estadística, Geografía e Informática (I.N.E.G.I.), es un auxiliar muy importante para los trabajos estadísticos ya que conserva mucha información importante que puede ayudar al estudio que se quiera realizar.

Si la información que se desea recabar no se encuentra en alguna dependencia, entonces es necesario realizar una investigación de campo para obtenerla y en este caso es necesario definir con precisión qué datos nos interesan para planear la forma de obtenerlos, ya que no solamente se estudian conjuntos de personas sino de cosas, objetos, entidades y desde varios puntos de vista; en las cuales será necesario medir características distintas.

Veamos el siguiente ejemplo:

Del conjunto de trabajadores de una industria, nos interesaría conocer varias características tales como:

- a) La edad para saber cuántos están próximos a su retiro o para determinar la edad promedio de los trabajadores.
- b) El peso promedio para saber si pueden desarrollar determinado trabajo donde el peso es importante.
- c) El grado de estudios de cada uno para saber si se les puede impartir cursos de preparación para el trabajo que requieren conocimientos de nivel medio superior, etcétera.

De los anterior podemos concluir que de acuerdo con la necesidad, se define la variable que se va a medir.

Una vez definida la variable se determina la población estadística.

Definición

Población estadística es un conjunto de personas, entidades u objetos del cual se quiere saber algo que nos interesa para tomar una determinación acertada.

Para facilitar el estudio de las poblaciones éstas se clasifican en:

- a) Población finita.
- b) Población infinita

Muestra

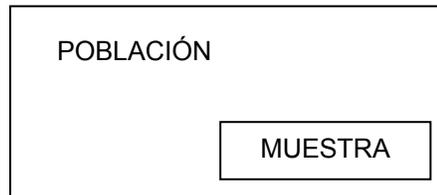
Una muestra es un conjunto de medidas u observaciones tomadas a partir de una población dada; es un subconjunto de la población. Desde luego, el número de observaciones en una muestra es menor que el número de posibles observaciones en la población, de otra forma, la muestra será la población misma. Las muestras se toman debido a que no es factible desde el punto de vista económico usar a toda la población. En algunos casos es imposible recolectar todas las posibles observaciones en la población.

Por ejemplo, si se desea estimar el gasto promedio anual de los estudiantes del C.B., se extraería una muestra formada por cierto número de estudiantes, se determinaría el gasto anual correspondiente a cada uno de ellos y después se obtendría el promedio. Se utiliza una muestra debido a que simplemente no se tiene el tiempo ni los recursos

para establecer el contacto con todos los estudiantes del C.B., aun cuando es posible hacerlo.

Definición

Muestra representativa es un subconjunto de la población que se estudia para determinar el parámetro que describe la característica deseada de la misma.



Todas las muestras son subconjuntos de la población pero no todas son representativas.

Las muestras representativas se selecciona aleatoriamente.

Definición

Muestra aleatoria es aquella que se obtiene de tal manera que cada posible observación disponible en la población, tiene la misma probabilidad de ser seleccionada.

Para poder obtener estas muestras es necesario que no intervenga la preferencia del investigador por algún elemento de la población; es decir, cada elemento de la población deberá tener igual oportunidad de ser seleccionado.

Los promedios y proporciones muestrales son características medibles de las muestras respectivas y se les llama estadísticas o estadígrafos.

Ejemplo:

En un campo de experimentación agrícola se ha desarrollado una variedad de jitomate. Se desea determinar el peso promedio de los jitomates de cada planta.

Del problema anterior determinar el parámetro y el estadígrafo.

Solución:

El parámetro de la población es: El peso promedio de todos los jitomates producidos por cada planta en una cosecha determinada.

ACTIVIDAD DE REGULACION

El estadístico o estadígrafo es el peso promedio de todos los jitomates producidos por planta, en una muestra aleatoria de plantas cultivadas de la cosecha.

Del siguiente problema identifica y escribe en la línea cuál es la población, la muestra, el parámetro y el estadígrafo.

De todos los estados de la República Mexicana se desea saber el ingreso bruto sobre recaudación de impuestos sobre la renta y el promedio de ingresos de diez de los estados tomados al azar.

La población es _____

La muestra es _____

El parámetro es _____

El estadígrafo o estadístico es _____

Los especialistas en Estadística emplean estadísticas muestrales para realizar inferencias acerca de los parámetros de la población.

El número de datos que constituye una muestra se llama tamaño de la muestra y se simboliza con la letra (n) . El número de datos de una población se simboliza con la letra (N) .

1.1.5 Clasificación de la Estadística

En base a lo que se ha dicho se concluye, que la Estadística como disciplina o área de estudio comprende técnicas descriptivas como inferenciales. Incluye la observación y tratamiento de datos numéricos y el empleo de los datos estadísticos con fines inferenciales.

Para su estudio se clasifica de la siguiente forma:

Estadística { Descriptiva
Inferencial

Definición

La Estadística es el estudio científico relativo al conjunto de métodos encaminados a la obtención, representación y análisis de observaciones numéricas, con el fin de describir la colección de datos obtenidos, así como inferir generalizaciones acerca de las características de todas las observaciones y tomar las decisiones más acertadas en el campo de su aplicación.

1.1.6 Estadística Descriptiva

El origen de la Estadística descriptiva puede relacionarse con el interés por mantener registros gubernamentales hacia fines de la Edad Media. Cuando los estados nacionalistas empezaron a surgir durante ese período, se volvió necesario obtener información acerca de los territorios bajo la jurisdicción de cada nación. Esta necesidad de información numérica acerca de los ciudadanos y recursos lleva al desarrollo de técnicos para obtener y organizar datos numéricos.

Hacia fines del siglo XVII, ya existían investigaciones semejantes a nuestros censos modernos. Al mismo tiempo, las compañías de seguros empezaban a recopilar tablas de mortalidad para determinar las primas de seguros de vida.

En las primeras etapas de desarrollo, la estadística incluía poco más que la obtención, clasificación y presentación de datos numéricos. Aún hoy en día, estas actividades siguen siendo una parte importante de la Estadística.

A continuación se da una definición de Estadística Descriptiva.

“La Estadística Descriptiva es el estudio que incluye la obtención, organización, presentación y descripción de información numérica”.

Ejemplo:

Un director de escuela desea conocer las aptitudes de cinco secretarías que trabajan en dicha institución.

Se aplica una prueba de aptitudes a las cinco secretarías y las calificaciones son 82, 85, 95, 92 y 91. La medida estadística que emplea el Director es la aptitud promedio o media aritmética, la cual es la suma de los valores obtenidos dividida por el número de observaciones. Entonces, la calificación promedio es:

$$\frac{82+85+95+92+91}{5} = \frac{445}{5} = 89$$

El cálculo de la media aritmética, simple como es, es una parte importante de la estadística descriptiva. El resultado se limita a los datos obtenidos en este caso particular y no implica ninguna inferencia o generalización acerca de las aptitudes de otras secretarías. Este método es de naturaleza descriptiva, debido a que el promedio condensa y describe la información obtenida, por ejemplo en el caso de las secretarías significa que el promedio de las aptitudes de las cinco secretarías es 89%.

La descripción de los datos también puede hacerse usando representaciones gráficas como veremos posteriormente.

1.1.7 Estadística Inferencial

Si el interés del Director de la escuela va más allá de la información obtenida, necesitará otras técnicas distintas a los métodos descriptivos.

Por ejemplo; podría desear conocer la aptitud promedio de las demás secretarías, pero carece del tiempo o de los recursos para aplicar una prueba a todas ellas. Podría utilizar la calificación promedio de las cinco secretarías como base para realizar una inferencia o estimación acerca de la aptitud promedio de todas las secretarías.

Con ese fin, necesitará conocer otra rama de la Estadística conocida como Estadística Inferencial o Inferencia Estadística.

Definición

“La inferencia estadística es una técnica mediante la cual se obtienen generalizaciones o se toman decisiones en base a una información parcial o completa obtenida mediante técnicas descriptivas”.

Para concluir diremos que existe otra gran división de las técnicas estadísticas:

- a) Estadística Paramétrica.
- b) Estadística No Paramétrica.

La Estadística Paramétrica es un conjunto de técnicas desarrolladas para niveles altos de medición como el de intervalos.

Los métodos paramétricos permiten hacer inferencias acerca de parámetros poblacionales de las distribuciones. Estos métodos fueron los primeros en ser desarrollados por los investigadores de la Estadística.

La Estadística no paramétrica es un conjunto de técnicas diseñadas para niveles de medición menores, por ejemplo, el nominal y ordinal, para efectuar estimaciones no habrá parámetros en estricto sentido.

A los procedimientos estadísticos que no dependen para su validez de la forma funcional de la distribución original de la población se les denomina procedimientos no paramétricos o libres de distribución.

Los Procedimientos No Paramétricos disponibles actualmente ofrecen varias ventajas para el investigador y analista de datos; entre ellos se pueden mencionar los que estableció Bradley en 1968:

- 1) La mayoría de los procedimientos no paramétricos se basan en un conjunto mínimo de suposiciones y esto tiende a reducir la posibilidad de utilizarlos inadecuadamente.
- 2) Los cálculos aritméticos necesarios para la aplicación de muchos procedimientos no paramétricos son cortos y fáciles, de manera que con su empleo se puede ahorrar tiempo.
- 3) Los procedimientos no paramétricos son por lo general fácilmente comprensibles para personas no muy formadas matemática o estadísticamente.
- 4) Se pueden aplicar los procedimientos no paramétricos cuando los datos que se van a analizar consisten más bien en rangos o conteos de frecuencia tales como porcentaje de pruebas, estatura, peso, longitud, entre otras.

1.2 CONCEPTOS BÁSICOS PARA EL ESTUDIO DE LA ESTADÍSTICA

El objetivo de la Estadística como disciplina, es ayudar a:

- a) Planear la búsqueda y obtención de la información.
- b) Organizar y sistematizar la información para su descripción y análisis.
- c) A partir de la información organizada, efectuar inferencias a través de la estimación y contrastación de hipótesis.

A través de estas etapas que estudiaremos en este fascículo, se obtiene la resolución de cualquier problema que nos interese. La Estadística nos proporciona los métodos correspondientes que facilitan la realización de cada una de las etapas que analizaremos.

a) PLANEACIÓN.

Para la planeación de cualquier actividad es necesario dar respuesta a los siguientes cuestionamientos:

1. ¿Qué problema me interesa resolver?
2. ¿Cómo lo voy a resolver?
3. ¿Para qué lo voy a resolver?

Dar respuesta a la primera pregunta implica determinar la población o muestra y fijar las variables para su estudio.

Como ya se dijo en párrafos anteriores, para disminuir tiempo y costo en lugar de trabajar con una población, se determina una muestra representativa de ésta, por lo que en la planeación se fija la forma de tomar la muestra.

Para dar respuesta al segundo cuestionamiento es necesario fijar la variable o variables de estudio, cómo se van a determinar, si ya existe esa información a quién debe solicitarse, si no existe entonces cómo se va a buscar, si es necesario una investigación de campo, elaborar el material necesario para realizarlo. La forma más usual para recolectar datos es mediante una encuesta, ésta debe elaborarse de tal forma que se tenga toda la información necesaria y evitar la innecesaria, que solamente dificulta el trabajo.

¿El para qué?

No permite tener presente el problema por resolver con el fin de fijar con precisión las variables que permitirán la resolución del mismo y con ello estar en posibilidad de tomar una decisión correcta.

La planeación se concluye con un breve curso informativo que se da los participantes en el trabajo de campo, con el fin de conocer los materiales que se deban llenar para recopilar la información.

Al finalizar el curso se procede a la recopilación de la información que se concentra para el trabajo de gabinete.

Las variables constituyen la herramienta fundamental de la Estadística, por que son la base esencial del estudio que se desea realizar y por tal motivo analizaremos cómo pueden ser éstas.

Las variables son:

- Características.
- Atributos.
- Rasgos.
- Cualidades.

ACTIVIDAD DE REGULACION

Investiga en un diccionario el significado de cada uno de los términos anteriores.

La variable es un propiedad intrínseca de individuos, objetos o grupos que interesan para un determinado estudio científico.

1. Ejemplos de variables:

- a) Un grupo de científicos desea conocer el I.Q. de un grupo de niños (variable "I.Q.").
- b) El director de una obra teatral desea conocer el color de ojos de un grupo seleccionado para el montaje (variable "color de ojos").
- c) Los cirujanos de una clínica desean conocer el tipo de nariz que tiene cada habitante de una colonia (variable "tipo de nariz").
- d) El director General del Colegio de Bachilleres desea conocer el índice de aprovechamiento de los alumnos de un plantel, en la materia de Estadística (variable "índice de aprovechamiento en Estadística").
- e) El gerente de una empresa desea conocer la dureza de ciertos lingotes de acero que adquirió en la planta de fundición (variable "dureza del acero").

En este ejemplo los habitantes se agrupan en tres categorías de acuerdo a su preferencia:

Los que prefieren la marca Ford.

Los que prefieren la marca Nissan.

Los que prefieren la marca Chrysler.

Cuando los valores de una variable permiten únicamente ubicar a cada individuo en una categoría y no hay orden entre estos valores la variable se llama "Cualitativa".

3. El maestro del grupo 502 del plantel 2 "Cien Metros", evaluó el grado de aprovechamiento en el curso de estadística, bajo la siguiente escala: Excelente, Bien, Regular, Mal; habiendo obtenido los siguientes resultados:

R, B, M, R, E, M, B, R, R, M, B, E, B, R, B,
B, R, B, B, R, B, M, E, R, R, B, B, E, B, R,
R, R, B, B, R, B, R, R, B, E, M, R, B, R,

Del problema anterior constes las siguientes preguntas:

¿Cuál es la variable que se está evaluando? _____

¿Cuántos alumnos obtuvieron E? _____

¿Cuántos alumnos obtuvieron B? _____

¿Cuántos alumnos obtuvieron R? _____

¿Cuántos alumnos obtuvieron M? _____

¿Cuántos elementos tiene la muestra? _____

¿En cuántas categorías se agrupan los elementos? _____

¿Qué tipo de variable es la que se está evaluando? _____

Cuando los valores de una variable cualitativa se pueden agrupar en un determinado orden, la variable se conoce como "variable cualitativa ordinal".

ACTIVIDAD DE REGULACION

En el ejemplo anterior indica otra forma de ordenar los resultados: _____

Define con tus propias palabras:

a) Variable cualitativa nominal: _____

b) Variable cualitativa ordinal: _____

c) Escribe en tu cuaderno 3 ejemplos donde la variable que se investiga sea: "Cualitativa nominal".

d) Escribe 3 ejemplos donde la variable que se investiga se: "Cualitativa ordinal".

Definición

Variable cualitativa es cuando solamente se busca en ella una cualidad o un atributo.

Variable cualitativa nominal es aquella que agrupa los elementos en categorías sin tener un orden.

Variable cualitativa ordinal es cuando las categorías en que se agrupan los elementos, pueden ser ordenados.

4. El médico de una guardería desea saber el crecimiento que tuvo cada niño a su cuidado, durante los primeros 6 meses del año, para ello se obtuvieron los siguientes resultados en centímetros:

8, 8, 7, 5, 4, 3, 4, 7, 5, 9, 3, 4, 7, 6, 5, 7, 3
5, 4, 5, 3, 9, 7, 6, 8, 4, 6, 9, 7, 8, 3, 4, 9, 5

Escribe la variable que se investiga en este ejemplo: _____

¿Qué tipo de variable es? _____

En este ejemplo puedes observar que a la variable (crecimiento en 6 meses) se le puede asociar un número que la define.

Si a la variable se puede asociar un número mediante el cual podemos hacer comparaciones u operaciones, esta variable se llama “Cuantitativa o numérica”.

ACTIVIDAD DE REGULACION

El director del hospital “1° de Octubre” del ISSSTE desea saber el número de pacientes atendidos en la sala de Urgencias, en el mes de septiembre, y para ello se obtuvieron los siguientes datos:

18, 25, 15, 30, 22, 27, 15, 18, 21, 19,
28, 27, 16, 19, 18, 30, 17, 16, 20, 15,
14, 19, 23, 27, 17, 14, 16, 20, 21, 28,

¿Cuál es la variable que se investiga? _____

¿Qué tipo de variable es? _____

Si solamente hay dos doctores en la sala, qué decisión tomarías si tu fueras el Director? ¿Por qué? _____

Definición

Variable cuantitativa es aquella que se puede asociar con un número con el cual podemos realizar operaciones o comparaciones.

En el ejemplo anterior, la variable es “pacientes atendidos en el día”, a ésta se le asocia un número que cuantifica a los pacientes atendidos. Esta variable nunca podrá tomar valores fraccionarios, por que siempre se hablará de personas y no de partes de una persona; en estos casos en que la variable siempre tomará valores enteros se le llama variable “discreta”.

Definición

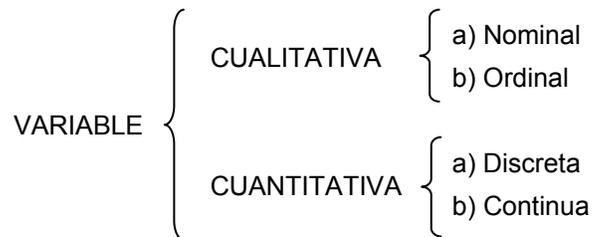
Variable cuantitativa discreta es la que siempre se asocia con valores enteros.

En el ejemplo número 4 donde la variable es “crecimiento durante 6 meses”, en los datos obtenidos podemos ver que la variable toma valores fraccionarios, en estos casos la variable se llama “Continua”.

definición

Variable cuantitativa continua es la que se puede asociar con valores de un intervalo de número reales.

Las variables estudiadas de acuerdo con sus características, se resumen en el siguiente cuadro:



ACTIVIDAD DE REGULACION

En el siguiente problema analiza los resultados y escribe en la línea el tipo de variable que se trate.

Los currícula de los investigadores de la UNAM registran los datos siguientes:

1. Nombre: _____
2. Edad: _____
3. Estatura: _____
4. Grado académico: _____
5. Número de investigaciones: _____
6. Experiencia en tipos de investigación: _____
7. Número de conferencias impartidas: _____

De los ejemplos anteriores habrás notado la importancia que tienen las variables en el estudio estadístico de un problema.

1.2.1 Presentación de Datos

Una vez que se han obtenido los datos y que se ha hecho el estudio de los valores que pueden tomar las variables, la primera tarea de la Estadística es la de ordenar y presentar los datos en tablas que permitan ver la tendencia de los mismos. Ordenados los datos se facilita su representación en diagramas y gráficas de diferentes tipos.

En esta unidad se verá la forma de describir, presentar, ordenar, resumir la información en tablas y su presentación en diferentes tipos de gráficas.

1.2.2 Distribución de Frecuencias

Los datos agrupados en tablas, nos permiten ver con facilidad el número de observaciones iguales o comprendidos en un intervalo, a este número de repeticiones iguales de la variable se llama frecuencia y se denota por f_i . Otros valores relacionados con la frecuencia son:

La frecuencia relativa que se denota por f_r .

La frecuencia acumulada que se denota por F_i .

La frecuencia relativa acumulada que se denota F_r .

En esta unidad analizaremos, ejemplificaremos y gratificaremos los datos de estos conceptos.

Escalas de Medición.

La clasificación que hemos hecho de las variables, depende del nivel de medición de la característica deseada.

En nivel de medición también denominados escalas de medición, lo podemos clasificar en:

1. Nominal.
2. Ordinal.
3. Por intervalo.

De acuerdo con esta clasificación podrás notar que dependen del tipo de variable que se analiza y por lo tanto reciben el mismo nombre de éstas.

5. En la carrera de caballos que se realizó el jueves 16 de septiembre corrieron 10 caballos, los cuales se numeraron de la siguiente forma:

10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20

Al finalizar la carrera, el primer lugar lo obtuvo el 13, el segundo lugar el 19, y el tercer lugar el 16. El primer lugar hizo un tiempo de 15.3 minutos, el segundo lugar 15.5 y el tercer lugar 15.8 minutos.

Definición de variables:

- a) Número de caballo (variable cualitativa nominal).
- b) Lugar que ocupó (variable cualitativa ordinal).
- c) Tiempo que hizo (variable cuantitativa continua).

Definición de escalas de medición:

- a) Nominal (numeración de los caballos).
- b) Ordinal (lugar ocupado en la carrera).
- c) Por intervalos (tiempo durante el recorrido).

Del ejemplo anterior podemos establecer las siguientes definiciones:

Escala nominal es la que se usa para asignarle una etiqueta a las categorías que se construyen de la variable con el único fin de distinguir unas de otras.

Escala ordinal es la que permite ordenar o jerarquizar las categorías que se construyen de la variable que se evalúa.

Escala de medición por intervalos es la que permite clasificar, ordenar y cuantificar las categorías que se establecen de la variable.

La escala de medición por intervalos es la de mayor nivel de medición e incluye las dos anteriores; para poder usarse con la precisión deseada es necesario fijar un patrón de medida que cuantifique a la variable con la misma exactitud, cuantas veces sea medida.

Algunos de los patrones que se usan son:

Años, kilos, litros, pesos, etcétera.

ACTIVIDAD DE REGULACION

Analiza el siguiente problema y establece todas las variables que consideres importantes, clasifícalas y determina el tipo de escala que se usaría en cada una.

El 16 de septiembre de un determinado año, se llevó a cabo el maratón de la libertad; en él participaron 15 mil maratonistas, quienes por su esfuerzo todos recibieron una medalla de participación y los premios que se repartieron fueron los siguientes:

1er. lugar: \$ 100,000.00

2do. lugar: una residencia

3er. lugar: un Gran Marqués

4to. lugar: un VW

5to. lugar: una motocicleta

1.2.3 Distribución de Frecuencias Absolutas y Relativas

Las primeras tareas de la Estadística descriptiva son ordenar, clasificar y resumir los datos obtenidos en la investigación de campo, para ello se concentran en tablas de frecuencia y éstas pueden ser:

- a) Absoluta.
- b) Relativa.
- c) Acumulada.

Con el análisis de las frecuencias podemos determinar la tendencia de la variable en estudio que como ya se dijo, ésta puede ser nominal, ordinal o cuantitativa y sus respectivas escalas de medición: nominal, ordinal o por intervalos, respectivamente.

6. Analicemos el siguiente ejemplo:

La maestra de orientación del Plantel 11 dio una conferencia al grupo 603 sobre las características y bondades de las carreras de Ingeniería, Química, Metalúrgica y Actuaría. Al final de la conferencia pidió que llenaran un cuestionario donde especificaron además de los datos personales, la carrera de preferencia. Se obtuvieron los siguientes resultados:

I, A, M, Q, Q, M, A, I, M, Q, A, Q, I, Q, M,
Q, M, M, A, Q, I, Q, M, I, I, Q, M, M, A, I,
M, A, A, Q, I, M, Q, Q, A, M, A, Q, M, A, Q,

En base a los datos del problema contesta lo siguiente:

- a) Establece la variable que se analiza. _____
- b) ¿Qué tipo de variable es? _____
- c) ¿Qué tipo de escala define la variable? _____

La forma en que se obtuvieron los datos, se muestra en la tabla anterior. Como podrás observar en esta forma es difícil interpretar la información, por lo que elaboramos la siguiente tabla de frecuencias:

- 1. Carrera que prefieren los alumnos del grupo 603 del Plantel 11 del Colegio de Bachilleres.

CATEGORÍAS	NÚMERO DE ALUMNOS	f
I	IIII III	8
Q	IIII IIIII	14
M	IIII IIII	13
A	IIII II	10
Total		45

- 7. Encuesta realizada por la maestra de orientación del Plantel 11, el 12 de septiembre de 1993.

El número de columnas de una tabla es variable y depende de la información que se quiera registrar.

En nuestro ejemplo podemos suprimir la columna 2 que representa el conteo de la variable el cual se puede realizar en otras hojas de trabajo.

En la tercera columna se registra la frecuencia.

definición

Frecuencia es el número de veces que se repite la misma observación. Se simboliza con f_i .

Analiza la tabla y contesta las siguientes preguntas:

- a) ¿Cuál categoría es la de mayor frecuencia? _____

b) ¿Qué información podemos determinar con la categoría de mayor frecuencia?

c) ¿Qué carrera es la menos solicitada? _____

De la tabla del ejemplo anterior puedes notar que las partes de una tabla de frecuencias debe contener las siguientes partes:

1. Título. Este describe la información más importante del problema como es:
 - a) La variable.
 - b) La muestra o población.
 - c) A quién corresponde la muestra.
2. Encabezado. Este describe el tipo de información que se describe en cada columna.
3. Cuerpo. El cuerpo agrupa el contenido de la información.
4. Final. En el final se registran los totales.
5. Fuente. En esta parte se debe especificar: cómo, quién, en dónde y cuándo se tomaron los datos.

Estas partes son comunes a todas las tablas que se elaboren en un estudio, para que el que las analice tenga toda la información y pueda hacer deducciones de los resultados.

Del ejemplo anterior se deduce la siguiente información:

- a) Variable: Carrera de preferencia.
- b) Tipo de variable: Cualitativa nominal.
- c) Tipo de escala: Nominal.
- d) Carrera de mayor aceptación: Química.
- e) Carrera de menor aceptación: Ingeniería.

8. Analicemos el siguiente ejemplo:

Se realiza un torneo relámpago de fútbol entre los planteles del norte del Colegio de Bachilleres que son el 1, 2, 6, 7, 9, 11, 12, 13, 18, 19. El torneo se realizó en tres fechas con sede en el Plantel de menor número. En la primera vuelta se sortean los

equipos pares e impares. Para la segunda ronda se sortean para jugar un par con un impar, y la tercera y última ronda se jugará de acuerdo con el lugar que tengan al final de la segunda ronda. Se formaron las parejas en el orden siguiente: 1er. lugar con 2do. lugar, 3º. con 4to., etcétera.

Los premios serán hasta el 5to. lugar y serán:

1er. lugar: Diploma y \$ 5,000.00

2do. lugar: Diploma y \$ 3,000.00

3er. lugar: Diploma y \$ 1,000.00

4to. lugar: Diploma y un paquete de libros clásicos para cada uno.

5to. lugar: Diploma.

Los resultados que se obtuvieron al finalizar el torneo son:

RESULTADOS DE CADA EQUIPO

Vuelta	Plantel	1	2	6	7	9	11	12	13	18	19
1	Goles	3	5	4	3	2	5	2	4	3	5
2	Goles	2	3	1	0	3	4	2	3	1	3
3	Goles	4	5	0	2	4	3	1	1	0.	2
Totales		9	13	5	5	9	12	5	8	4	10

Con los datos anteriores formamos la tabla de frecuencias:

1. Resultados obtenidos del torneo relámpago efectuado por los 10 planteles de la zona norte, efectuado los tres primeros domingos del mes de septiembre de 1993.

Plantel	Número de goles (f_i)	Lugar que ocupó
2	13	1º.
11	12	2º.
19	10	3º.
1	9	4º.
9	9	4º.
13	8	5º.
6	5	6º.
7	5	6º.
12	5	6º.
18	4	7º.
Total	80	

- 4.
5. $T_{\bar{x}}$

torneo relámpago.

el

Del ejemplo anterior contesta las siguientes preguntas:

- a) ¿En cuántas categorías se agrupan los elementos del problema?
- b) ¿Cuál es la variable del problema?
- c) ¿Qué valores toma la variable?

Analicemos el siguiente problema:

9. El gerente de una Empresa, Kimberly preocupado por el pago de energía consumida solicitó al jefe de planta, un estudio del consumo diario durante el mes de agosto. Los resultados obtenidos. KW/hr (kilowatts por hora) son los siguientes:

Con esta información ordenamos los datos de una tabla de frecuencias:

1. Consumo de energía en KW/hr de la empresa Kimberly correspondiente al mes de agosto de 1993.

2.

Consumo KW/hr	f_i
5	2
6	3
7	4
8	5
9	6
10	5
11	3
12	2
Total	30

3.

4.

5. Investigación elaborada por el jefe de planta, datos tomados del consumo diario del mes de agosto de 1993.

Del problema anterior contesta las siguientes preguntas:

- a) Define la variable del problema: _____
- b) ¿Qué tipo de variable es? _____
- c) ¿Qué valores toma la variable? _____

- d) ¿Qué tipo de escala define la variable? _____
- e) ¿Cuál es la mayor frecuencia de la variable? _____
- f) ¿Qué frecuencia tiene la variable cuya categoría es 10? _____

Respuestas:

- a) La variable del problema es “consumo en KW/hr”.
- b) La variable es de tipo cuantitativo o numérica continua.
- c) Los valores de la variable se encuentran en un rango de 5 a 12 KW/hr.
- d) El tipo de escala que define la variable es por intervalos.
- e) La mayor frecuencia es 6 y significa en 6 días del mes se tiene consumo de 9 KW/hr.
- f) La frecuencia de la categoría 10 es 5 lo cual significa que en 5 días del mes hubo un consumo de 10 KW/hr.

Ya quedó establecido que el número de veces que se repite la misma observación se llama frecuencia (f_i) o frecuencia absoluta y el conocimiento de esta variable nos permite inferir otro conocimiento. En nuestro ejemplo podemos ver en cuántos días hubo el mismo consumo, en cuántos días hubo el menor consumo o en cuántos días hubo el mayor consumo.

También podemos inferir hacia dónde se carga el mayor o menor consumo.

Otro parámetro importante es la frecuencia relativa que simbolizaremos con “ f_r ”, ésta se obtiene dividiendo la frecuencia absoluta (f_i) entre el número de elementos de la muestra que simbolizaremos con (n).

La definición matemática es:

$$f_r = \frac{f}{n} \dots \dots \dots (1)$$

De la ecuación (1) puedes observar que la frecuencia relativa se puede expresar como una razón, como una proporción o como un porcentaje (%).

Veamos el siguiente ejemplo:

- 10. La puntuación obtenida en un examen que se aplicó a 100 obreros de la fábrica de vidrio el Fanal, es la que se muestra en la siguiente tabla de frecuencias:

1. Resultados del examen aplicado a 100 obreros de la fábrica de vidrio el Fanal.

2.

Puntuación	f_i	Razón	Frecuencia relativa (f_r)	
			Proporción	Porcentaje
0	1	1/100	0.01	1 %
1	2	2/100	0.02	2 %
2	3	3/100	0.03	3 %
3	5	5/100	0.05	5 %
4.	6	6/100	0.06	6 %
5	8	8/100	0.08	8 %
6	20	20/100	0.20	20 %
7	25	25/100	0.25	25 %
8	15	15/100	0.15	15 %
9	10	10/100	0.10	10 %
10	5	5/100	0.05	5 %
Total	100	100/100	1.00	100 %

5. Investigación realizada por el jefe del departamento de capacitación de la fábrica de vidrio el Fanal, el 5 de septiembre de 1993.

Analiza la tabla anterior y contesta las siguientes preguntas:

- ¿Cuál es la variable del problema? _____
- ¿Qué escala define a la variable? _____
- ¿Qué puntuación tiene la mayor frecuencia? _____
- ¿Qué porcentaje de obreros reprobó el examen si la calificación aprobatoria es de 6 a 10? _____
- De este resultado, ¿qué puede inferir el jefe del departamento de capacitación? _____

Definición

Frecuencia relativa (f_r) es la proporción de elementos que pertenecen a una categoría y ésta se obtiene dividiendo su frecuencia absoluta entre el número total de elementos de la muestra.

ACTIVIDAD DE REGULACION

Elabora una tabla de frecuencias de los ejemplos 4, 5, 6 y agrega en la tabla la columna de frecuencia relativa (f_r) en sus tres formas de expresión.

Hasta el momento, en los problemas que hemos analizado las muestras son pequeñas (n es pequeño). Sin embargo cuando la muestra o población se compone de un considerable número de elementos, la tabla de frecuencias se elabora agrupando los datos en clases y ahora la tabla se llama "Tabla de frecuencias con datos agrupados".

La formación de clases o intervalos de clase que simbolizaremos con (k) es muy variado y depende generalmente del tamaño del rango de la población o muestra.

Definición

Rango (R) es el intervalo en que se distribuyen los datos en observaciones de una muestra y se determina restándole al mayor valor el menor valor.

La definición matemática del rango es:

$$R = X_n - x_1 \dots\dots\dots (2)$$

X_n = valor mayor

X_1 = valor menor

No existe alguna ley que defina cómo obtener el número de clases; pero la experiencia recomienda que no sea menor que 5 ni mayor de 20, esto es:

$$5 \leq k \leq 20 \dots\dots\dots (3)$$

k = número de clases

Una vez definido el número de clases (k), para obtener la amplitud de clase (A) aplicamos la siguiente ecuación:

$$A = \frac{R}{K} \dots\dots\dots (4)$$

Otra forma de determinar el número de intervalos de clase (k) es mediante la ecuación de Sturges y ésta es:

$$K = 1 + 3.322 (\log n) \dots\dots (5)$$

Donde:

k = número de intervalos
n = tamaño de la muestra
log = logaritmo en base 10

Veamos el siguiente ejemplo:

11. El gerente de una compañía de ventas al mayoreo de diferentes tipos de mercancías desea conocer el comportamiento de las llamadas telefónicas durante los meses de marzo y abril del año en curso; por lo que le encomienda a su secretaria que realice esa investigación. La secretaria obtuvo los siguientes datos, en número de llamadas por día:

30, 38, 36, 35, 29, 28, 30, 35, 40, 48, 50, 20, 25, 56, 30
27, 29, 46, 41, 31, 31, 31, 39, 28, 36, 37, 52, 44, 49, 52
56, 58, 40, 39, 38, 40, 27, 24, 30, 32, 35, 38, 26, 25, 24
60, 55, 48, 37, 31, 30, 22, 20, 24, 26, 23, 22, 28, 27, 48

La secuencia de operaciones es la siguiente:

1. Ordenamos la información en sentido creciente.
2. Determinamos X_1 y X_n .
3. Calculamos el rango R.
4. Calculamos k mediante la ecuación de Sturges.
5. Determinamos la amplitud de la clase A.
6. Elaboramos la tabla de frecuencias con datos agrupados; para ello colocamos el primer intervalo en el primer renglón y formamos los siguientes de acuerdo con la amplitud.

Respuestas:

1. 20, 20, 22, 22, 23, 24, 24, 24, 25, 25, 26, 26, 27, 27, 27, 28
28, 29, 29, 30, 30, 30, 30, 30, 31, 31, 31, 32, 35, 35, 35, 36
36, 37, 37, 38, 38, 38, 38, 39, 39, 40, 40, 40, 41, 44, 46, 48
48, 48, 49, 50, 52, 52, 55, 56, 56, 57, 58, 60
2. $X_1 = 20$
 $X_n = 60$
3. $R = X_n - X_1 = 60 - 20 = 40$
4. $K = 1 + 3.32 \ln(60) = 1 + 3.322(1.7781) = 1 + 5.9 = 7$

$$5. A = \frac{R}{K} = \frac{40}{7} = 5.7$$

Cada uno de los intervalos de clase debe contener 6 valores en total.

6. Determinamos cada intervalo:

1. Número de llamadas telefónicas correspondientes a los meses de marzo y abril.

2.	Clases (K)	f	Razón	Frecuencia relativa (f_r)	
				Proporción	Porcentaje
	20-25	10	10/60	0.17	17%
	26-31	17	17/60	0.28	28%
	32-37	8	8/60	0.13	13%
3.	38-43	10	10/60	0.17	17%
	44-49	6	6/60	0.10	10%
	50-55	4	4/60	0.07	7%
	56-61	5	5/60	0.08	8%
4.	Total	60	60/60	1.00	100 %

5. Información investigada y elaborada por la secretaria de la gerencia de la empresa.

El uso de los intervalos de clase es con la finalidad de condensar la información para facilitar su manejo.

Los valores extremos de cada clase se conocen como límites de clase. El valor menor se llama límite inferior de la clase y el mayor se llama límite superior de clase.

Analiza la tabla y contesta las siguientes preguntas:

- a) Escribe el límite inferior de la clase de mayor frecuencia _____
- b) Escribe el límite superior de la clase de menor frecuencia _____
- c) ¿Qué porcentaje representa la clase de mayor frecuencia? _____

En el ejemplo anterior la variable es discreta porque las llamadas telefónicas siempre se toman como valores enteros toda vez que no podemos hablar de fracciones de llamadas telefónicas.

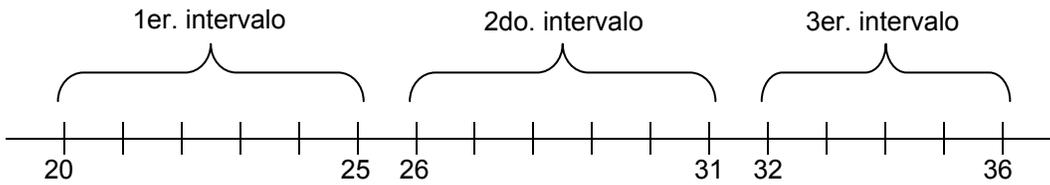
Si analizamos los intervalos de clase del ejemplo anterior podemos notar lo siguiente:

De un intervalo a otro hay un salto por ejemplo:

1er. intervalo 20 – 25

2do. Intervalo 26 – 31

Veámoslo en una gráfica lineal.



En la gráfica podemos ver que al formar los intervalos de clase hay un valor entre clase y clase que se pierde.

Como la variable es discreta sabemos que entre estos valores no hay ninguna información que se pierda; pero ¿qué pasa si la variable es continua?, en estos casos si hay la posibilidad que entre el 25 y 26 se pierdan los valores comprendidos como es 25.1, 25.3, 25.6, etcétera.

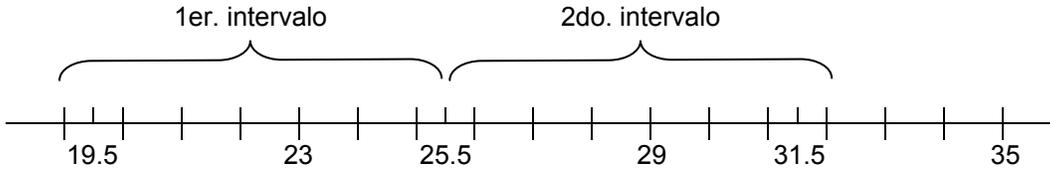
Para evitar este error, si la variable es continua, entonces después de haber determinado los límites de clase, se fijan otros límites que inician medio punto antes y medio punto después; de esta forma no hay posibilidad de perder información.

A cada uno de estos nuevos límites se le llama “límite real de clase”.

Veamos el mismo ejemplo anterior.

Límites de clase	Límites reales de clase
CLASES	LÍMITES REALES
20-25	19.5 – 25.5
26-31	25.5 – 31.5
32-37	31.5 – 37.5
38-43	37.5 – 43.5
44-49	43.5 – 49.5
50-55	49.5 – 55.5
56-61	55.5 – 61.5

Si representamos los límites reales de clase en una gráfica lineal podemos observar que ya no hay saltos entre cada clase:



De la tabla anterior contesta las siguientes preguntas:

- a) Escribe en la línea el límite real inferior de la tercera clase _____
- b) Escribe en la línea el límite real superior de la quinta clase _____
- c) Escribe en la línea el límite superior de la segunda clase _____

Al cambiar los límites reales de clase, el valor del intervalo de clase (A) no cambia y para determinarlo aplicamos la siguiente ecuación:

$$A = X_s - X_i \dots\dots\dots (6)$$

X_s = Límite superior de cualquier clase

X_i = Límite inferior de la misma clase considerada

Por ejemplo para la 4ta. Clase del ejemplo anterior

$X_s = 43.5$ Sustituyendo en (6) obtenemos:

$X_i = 37.5$ $A = 43.5 - 37.5 = 6$

Apliquemos la misma ecuación para la misma clase si la variable es discreta:

$X_s = 43$ $A = 43 - 38 = 5$

$X_i = 38$

De este resultado se infiere que para la variable discreta debe aplicarse la siguiente ecuación:

$$A = X_s - X_i + 1 \dots\dots\dots (7)$$

$$A = 43 + 38 + 1 = 6$$

Otra característica importante del intervalo de clase o marca de clase es el punto medio de clase o marca de clase (M_i).

Definición

La marca de clase es el valor del punto que se localiza a la mitad del intervalo de cada clase o intervalo real de clase.

Su definición matemática es:

$$M_i = \frac{X_s - X_1}{2} \dots\dots\dots (8)$$

Determinemos los puntos medios o marcas de clase del mismo ejemplo anterior en los dos tipos de intervalos.

Límites de clase (variable discreta)		Límites reales de clase (variable continua)	
CLASES	M_i	CLASES	M_i
20-25	22.5	19.5 – 25.5	22.5
36-31	28.5	25.5 – 31.5	28.5
32-37	34.5	31.5 – 37.5	34.5
38-43	40.5	37.5 – 43.5	40.5
44-49	46.5	43.5 – 49.5	46.5
50-55	52.5	49.5 – 55.5	52.5
56-61	58.5	55.5 – 61.5	58.5

De la tabla se deduce que el punto medio de clase (M_i) o marca de clase es el mismo en cada clase para ambas variables (discreta o continua).

ACTIVIDAD DE REGULACION

En el siguiente problema, elabora una tabla de frecuencia de datos agrupados con 5 columnas con el siguiente encabezado:

1. Clases.
2. Límites reales de clase.
3. Punto medio de clase o marca de clase.
4. Frecuencia absoluta (f_i).
5. Frecuencia relativa (f_r).

Problema: En una cadena de tiendas, la venta de líquido anticongelante para automóviles, (en número de litros) en un determinado día de diciembre fue el que se indica:

216, 269, 235, 235, 224, 249, 244, 240, 252, 210, 256
 228, 233, 239, 253, 237, 221, 273, 229, 236, 215, 260
 230, 228, 230, 255, 245, 244, 230, 231, 240, 222, 223

Hasta el problema anterior hemos mostrado cómo organizar los datos en una tabla de frecuencia de datos agrupados en clases y también hemos incluido la fracción o porcentaje de cada clase (f_r). Esta información nos ha permitido hacer algunas inferencias.

A continuación veremos otra característica de la muestra que nos permitirá un análisis más amplio de los resultados.

1.2.4 Distribución de Frecuencias Acumuladas.

La frecuencia acumulada (F_i) es otra característica de la muestra que nos permitirá determinar la posición de un caso particular que nos interese en comparación con el total de los elementos.

DEFINICIÓN:

Frecuencia acumulada (F_i) de una clase es la que se obtiene sumando las frecuencias de las clases anteriores con la frecuencia de ésta.

Su definición matemática es:

$$F_3 = \sum_{i=1}^3 f_1 + f_2 + f_3$$

$$F_n = \sum_{i=1}^n f_1 + f_2 + \dots + f_n \quad \dots \dots \dots (9)$$

Al calcular la frecuencia acumulada (F_i) podemos determinar su frecuencia relativa acumulada (F_r) en la forma ya explicada mediante la ecuación (1), esto es:

$$F_r = \frac{F}{n} \quad \dots \dots \dots (1)$$

Regresemos al problema (11) de las llamadas telefónicas y calculemos la frecuencia acumulada (f_1) y la frecuencia relativa acumulada (F_r).

1	2	3	4	5	6	7
CLASES	LÍMITES REALES	MARCA DE CLASE (M_i)	FREC. ABS. (f_i)	FREC. REL. (F_i)	FREC. ACUMULADA	
					F_i	F_R
20-25	19.5-25.5	22.5	10	0.17	10	0.17
26-31	25.5-31.5	28.5	17	0.28	27	0.45
32-37	31.5-37.5	34.5	8	0.13	35	0.58
38-43	37.5-43.5	40.5	10	0.17	45	0.75
44-49	43.5-49.5	46.5	6	0.10	51	0.85
50-55	49.5-55.5	52.5	4	0.07	55	0.92
56-61	55.5-61.5	58.5	5	0.08	60	1.00
Total			60	1.00		

La frecuencia acumulada para la 4ta. Clase es $F = 45$; de este valor se infiere que hasta esta clase corresponden 45 de las 60 observaciones realizadas.

También se infiere que a esta clase corresponden un número menor o igual a 43 llamadas telefónicas.

La frecuencia relativa de esta clase es $F = 0.75$. este valor significa que hasta esta clase corresponde el 75% de todas las llamadas.

Analiza la tabla y contesta las siguientes preguntas:

- ¿A qué clase corresponde F_2 y cuál es su valor? _____
- Indica el porcentaje de llamadas para F_5 _____
- ¿Qué número de llamadas telefónicas corresponde a F_5 ? _____

Como pudiste notar en la tabla anterior, la frecuencia acumulada nos permite determinar el número de llamadas menor o igual que el límite superior de la clase correspondiente, por lo tanto a esta columna la podemos significar con el término menor que.

Si queremos determinar el número de llamadas mayor que, lo que tenemos que hacer es desacumular la frecuencia y para ello en lugar de sumar restamos al número de observaciones (n) la frecuencia de la clase (f_i) correspondiente.

Veámoslo en la tabla del ejemplo anterior.

1	2	3	4	5	6	7	8
CLASES	MARCA DE CLASE (M_i)	FREC. ABS. (f_i)	FREC. REL. (f_i)	F _i ACUMULADA		FREC. RELAT.	
				menos que	más que	menos	más
20-25	22.5	10	0.17	10	50	0.17	.83
26-31	28.5	17	0.28	27	33	0.45	.55
32-37	34.5	8	0.13	35	25	0.58	.42
38-43	40.5	10	0.17	45	15	0.75	.25
44-49	46.5	6	0.10	51	9	0.85	.15
50-55	52.5	4	0.07	55	5	0.92	.08
56-61	58.5	5	0.08	60	0	1.00	.00
Total		60	1.00				

La columna 7 correspondiente a “más que” se interpreta de la siguiente forma: para $F = 15$ significa que 15 de los 60 días que se analizan, tuvieron un número de llamadas mayor o igual a 44 y a este número de llamadas equivale al 25% del total.

Analiza la gráfica y contesta las siguientes preguntas:

- a) Para $F = 25$ indica el número de llamadas que le corresponden y a qué porcentaje pertenecen.

ACTIVIDAD DE REGULACION

En el siguiente problema elabora las 10 columnas de la tabla de frecuencias y explica lo que se deduce de las clases 3 y 5.

En la caseta de cobro de la autopista a Pachuca ingresaron durante las 48 horas del 19 de septiembre, después del puente, el número de vehículos por hora que a continuación se indica:

13 19 22 14 13 16 19 21
23 11 27 25 17 17 13 20
23 17 26 20 24 15 20 21
10 22 18 25 16 23 19 20
21 17 18 24 21 20 19 26

Con los ejemplos anteriores pudiste notar que la tabla de frecuencias nos permite inferir ciertos conocimientos de la variable que se analiza.

Otra forma de analizar la información e inferir ciertos conocimientos, es mediante la representación gráfica de los mismos.

1.2.5 Gráficas

Al representar en una gráfica la información concentrada en la tabla de frecuencias, ésta es un recurso visual que nos permite tener una idea clara, precisa, global y rápida acerca de las observaciones de una muestra o población.

Existen muchos tipos de gráficas en las que se pueden representar la frecuencia absoluta (f_i), relativa (f_r) y acumulada (F_i) y con ellas podemos estimar algunos valores con la simple observación.

Los diferentes tipos de gráfica que podemos usar para representar las observaciones de un determinado problema y la selección de este tipo, dependen de la variable en estudio.

Si la variable en estudio es del tipo cualitativo, los gráficos pueden ser:

- a) De barras; horizontales o verticales.
- b) Circulares.
- c) Pictogramas, etcétera.

Si la variable en estudio es de tipo cuantitativo, los gráficos que podemos usar para su representación gráfica son:

- a) Histogramas.
- b) Polígonos de frecuencias que a continuación analizaremos.

1.2.6 Histograma y Polígono de Frecuencias.

El histograma es la forma más usual para analizar las características observables de una variable continua.

Definición

Histograma es la representación gráfica en el plano coordenado de las características concentradas en la tabla de frecuencias de una variable continua.

Para trazar el histograma, la secuencia de operaciones es:

1. En los ejes coordenados del plano cartesiano representamos los datos de la siguiente forma:
 - a) En el eje de las abscisas (horizontal) se representan las clases con sus límites reales de clase y las marcas de clase (M_i) de cada intervalo.
 - b) En el eje de las ordenadas (vertical) representamos las frecuencias absolutas en que ocurre la variable.

2. Por los límites reales superior e inferior de cada clase se trazan barras verticales que se cortan mediante una horizontal que se traza a la altura del punto correspondiente a la frecuencia de cada clase.
3. Por la naturaleza continua de la variable, los rectángulos se trazan adyacentes, toda vez que en esta forma se debe dividir el eje horizontal.
4. El área representada por cada barra es equivalente a la proporción de la frecuencia del intervalo de clase correspondiente con respecto al total.

Analícemos el siguiente problema:

11. Al gerente general de la empresa “Conductores Monterrey” le interesa conocer la antigüedad de sus trabajadores, por lo que le indica al gerente de personal que realice un análisis del problema.

El gerente de personal recabó de los expedientes la siguiente información sobre los años de antigüedad:

13, 19, 22, 14, 13, 16, 19, 21
 23, 11, 27, 25, 17, 17, 13, 20
 23, 17, 26, 20, 24, 15, 20, 21
 23, 17, 29, 17, 19, 14, 20, 20
 10, 22, 18, 25, 16, 23, 19, 20
 21, 17, 18, 24, 21, 20, 19, 26

Con esta información decidió representarlos en una gráfica (histograma). Recuerda la secuencia de operaciones que establecimos:

1. Ordenamos los datos en sentido creciente:

10, 11, 13, 13, 13, 14, 14, 15, 16, 16, 17, 17, 17, 17, 17, 17, 18,
 18, 19, 19, 19, 19, 19, 20, 20, 20, 20, 20, 20, 20, 21, 21, 21, 21,
 22, 22, 23, 23, 23, 23, 24, 24, 25, 25, 26, 26, 27, 29.

2. Calculamos el rango R, para ello determinamos los valores mayor y menor de las puntuaciones.

$$X_n = 29$$

$$X_1 = 10$$

3. Calculamos $R = X_n - X_1 = 29 - 10 = 19$

$$R = 19$$

4. Calculamos el número de clases (K), para ello determinamos (n)

$$N = 48; K = 1 + 3.322 \log 48 = 1 + 3.322 (1.68) = 1 + 5.58 = 6.58 \quad K = 7$$

5. Determinamos la amplitud de cada clase (A)

$$A = \frac{R}{K} = \frac{19}{7} = 2.7 \quad \therefore \quad A = 3$$

Se han redondeado los valores de K y A porque el número de clases y la amplitud de la clase nunca serán fraccionarios.

6. Determinamos cada intervalo de clase y para ello calculamos los límites de clase y los registramos en la primera columna de la tabla.

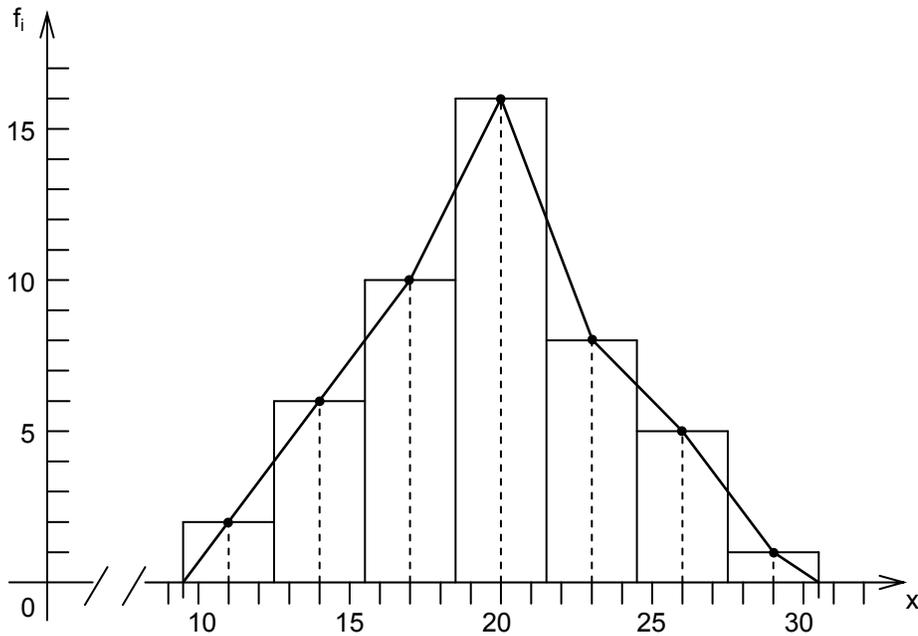
1. Números de años cumplidos de antigüedad de los obreros de la empresa "Conductores Monterrey", hasta el 30/IX/1993.

CLASES	LÍMITES REALES	P. MEDIO DE CLASE (M _i)	FREC. (f _i)	FREC. (f _r)	FRECUENCIA ACUMULADA		
					F MENOR Q	F30 MAYOR Q	F _r
10-12	9.5-12.5	11	2	0.042	2	46	0.042
13-15	12.5-15.5	14	6	0.125	8	40	0.167
16-18	15.5-18.5	17	10	0.208	18	30	0.375
19-21	18.5-21.5	20	16	0.333	34	14	0.708
22-24	21.5-24.5	23	8	0.167	42	6	0.875
25-27	24.5-27.5	26	5	0.104	47	1	0.979
28-30	27.5-30.5	29	1	0.021	48	0	1.00
Total			48	1.000			

5. Información investigada por el gerente de personal de la empresa.
6. Trazamos los ejes del plano coordenado, fijamos una escala para cada eje y representamos en el vertical las frecuencias y en el eje horizontal las clases.

La mayor frecuencia es $f_4 = 16$ por lo que con la escala establecida en cm. Marcamos 16 divisiones en el eje vertical.

En el eje horizontal no es necesario iniciar por el cero, en nuestro ejemplo podemos iniciar a partir de 9, indicando que se trunca una parte del eje horizontal.



Gráfica No. 1

Analiza la gráfica anterior y piensa que eres el gerente general. Explica con tus propias palabras toda la información sobre la antigüedad de los obreros, que puedas inferir de ella.

Comenta tus resultados con tu profesor o asesor.

Las líneas verticales punteadas corresponden al punto medio o marca de clase y éste nos indica el promedio de las puntuaciones en cada clase.

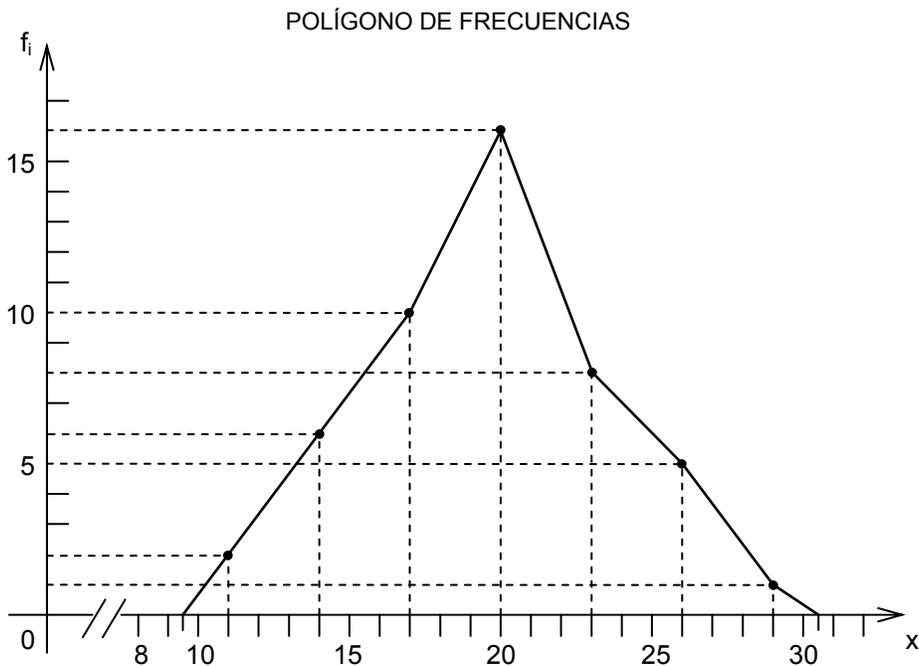
En nuestro ejemplo para la tercera barra, el promedio en años cumplidos de los obreros de la empresa ubicados en esa clase es de $M_3 = 17$ años y como la frecuencia es $f_3 = 10$ Obreros con la misma antigüedad.

Otra gráfica que permite describir los datos de una distribución de frecuencias es el polígono de frecuencias.

1.2.7 Polígono de Frecuencia

El polígono de frecuencia se construye a partir de los datos de la tabla de frecuencias. Sobre el eje horizontal se levanta por el punto medio segmentos verticales punteados que terminan a la altura de su frecuencia de clase, se unen los puntos superiores con un segmento de recta que empieza medio punto antes del límite superior de la última clase.

De nuestro ejemplo anterior obtenemos:



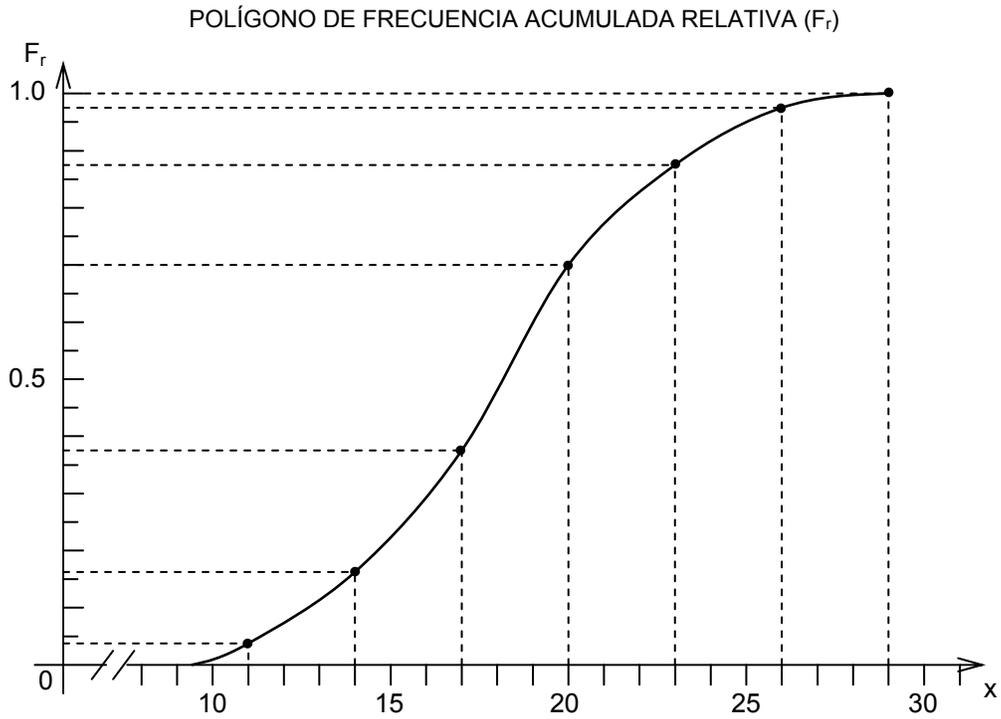
Gráfica No. 2

Otra forma de trazar el polígono de frecuencias es:

- Sobre el histograma se trazan segmentos punteados que inician en el punto medio de clase y terminan a la altura del rectángulo.
- Se unen los puntos finales de éstos segmentos con una línea continua que inicia medio punto antes del límite inferior de la primera clase y termina medio punto después del límite superior de la última clase (ver Gráfica No. 1).

Si en la escala vertical en lugar de representar las frecuencias absolutas, representamos la frecuencia relativa, obtenemos un polígono de frecuencia acumulada.

Veámoslo en el mismo ejemplo anterior.

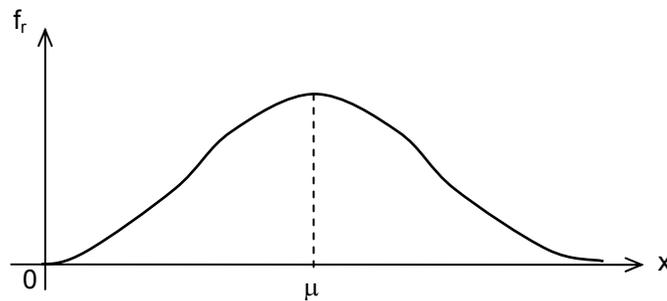


Gráfica No. 3

Analiza la gráfica e indica toda la información que puedes obtener.

Si observas las gráficas 2 y 3 harás notado que son muy parecidas y su tamaño dependerá de la escala que se use.

Si los intervalos de clase se toman cada vez más pequeños, entonces los puntos se unen con segmentos curvos que van suavizando la forma del polígono de frecuencias hasta tomar una forma parecida a la siguiente gráfica.



Gráfica No. 4

Al graficar las columnas 6 y 7 de la tabla de frecuencias de nuestro ejemplo, correspondientes a la frecuencia acumulada, obtenemos otras gráficas diferentes a las anteriores.

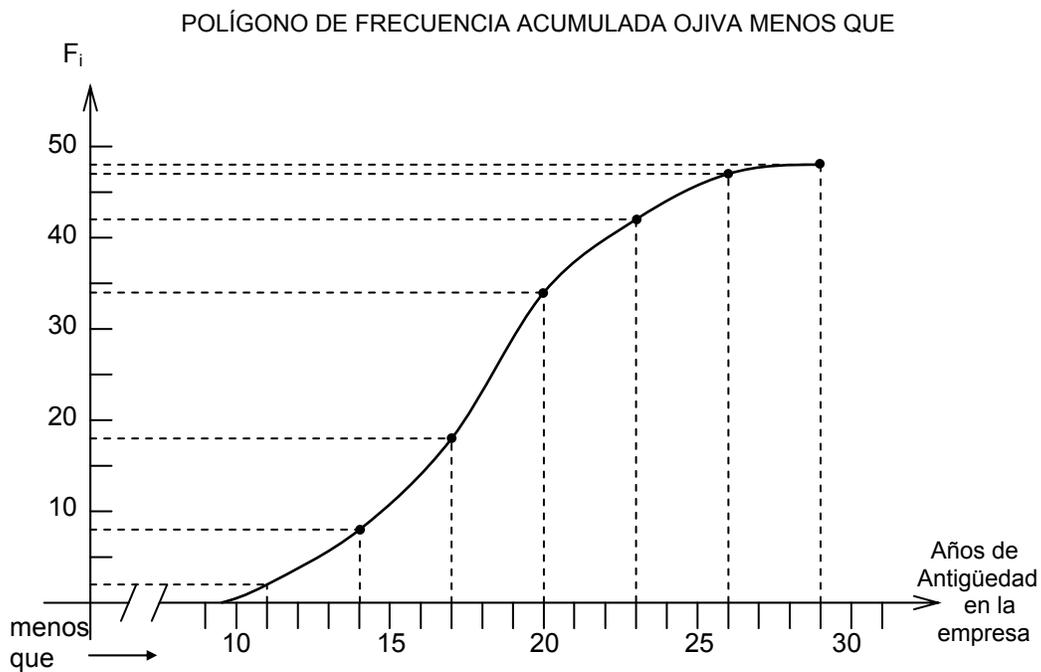
1.2.8 Polígono de Frecuencia Acumulada

La gráfica de la frecuencia acumulada es muy útil porque en ella podemos determinar cuántas observaciones hay por arriba o por debajo de algún valor que nos interese.

La gráfica que se obtiene de la frecuencia acumulada también se conoce con el nombre de ojiva.

Para trazar dicha gráfica se procede como en los gráficos anteriores, es decir, en el eje horizontal se trazan los intervalos de clase y marcas de clase y en el vertical las frecuencias acumuladas.

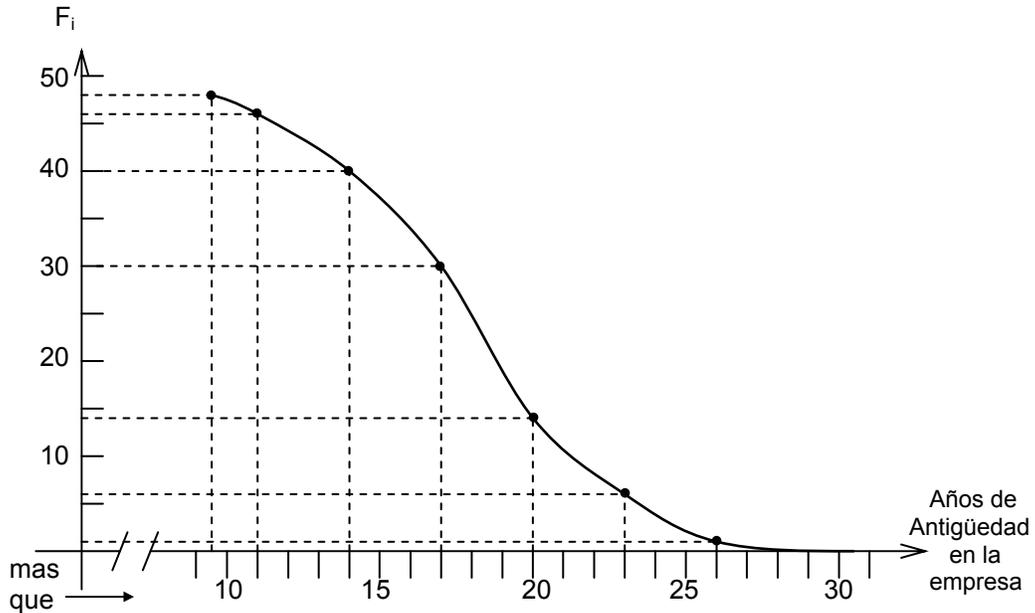
De nuestro ejemplo tomando los datos de la tabla y las columnas 6 y 7 obtenemos:



Gráfica No. 5

Analiza la gráfica y escribe con tus propias palabras la información que puedes deducir de ella.

Ya vimos que al desacomular las frecuencias obtenemos la gráfica mayor que, tomamos los datos de la columna 7 y obtenemos:



Gráfica No. 6

Analiza la gráfica y escribe en la línea el número de trabajadores que tengan una antigüedad mayor que:

- | | | |
|------------|------------|------------|
| a) 10 años | c) 18 años | e) 25 años |
| b) 13 años | d) 20 años | f) 30 años |

La variable del problema que hemos analizado es de tipo cuantitativa continua, si la variable que analices es cuantitativa discreta, entonces puedes darle el mismo tratamiento para el trazo de las gráficas o bien puedes prescindir de la columna número 2 y trazar el histograma con los saltos que da la variable de un intervalo a otro.

ACTIVIDAD DE REGULACION

Traza la gráfica de la columna 8 de nuestro ejemplo (f_i) y analiza la información que se puede determinar de ella.

Ahora analizaremos las gráficas de problemas cuya variable sea cualitativa.

1.2.9 Gráficas Circulares, Diagramas de Barras, Pictogramas y Gráficas de Líneas

Gráficas Circulares.

Una forma de representar datos u observaciones de una variable cualitativa es mediante un diagrama circular.

Para trazar la gráfica, se hace una distribución proporcional de las frecuencias del problema con respecto a la circunferencia determinando sectores circulares para cada categoría.

Veamos el siguiente ejemplo:

12. El gobierno de la ciudad de México promocionó cursos dominicales para todas las personas que quieran aprender alguna habilidad artesanal o artística; los cursos se realizarán de 8:00 a 12:00 hrs. Las inscripciones quedan abiertas todo el mes de septiembre.

Al final del mes, la inscripción fue la que se muestra en la siguiente tabla:

CURSOS	No. DE PARTICIPANTES (f_r)
Música	250
Pintura	350
Danza	200
Tejido	400
Bordado	100
Ajedrez	500
Total	1800

Para distribuir cada clase en la circunferencia se aplica una regla de tres simple relacionando el total de inscritos (1800) a 360° que tiene la circunferencia y de esta relación determinamos la parte que le corresponde a cada curso.

El planteamiento de cada distribución es:

$$\begin{array}{l} 1800 \text{ es a } 360^\circ \\ \text{como } 250 \text{ es a } x \\ 1800 \text{ ————— } 360^\circ \\ 250 \text{ ————— } x \end{array}$$

Recuerda que para hallar el resultado se multiplica como se indica con las flechas y se despeja la incógnita, obteniendo:

$$1800x = 360^\circ (250)$$

$$x = \frac{360^\circ (2500)}{1800} = 50^\circ$$

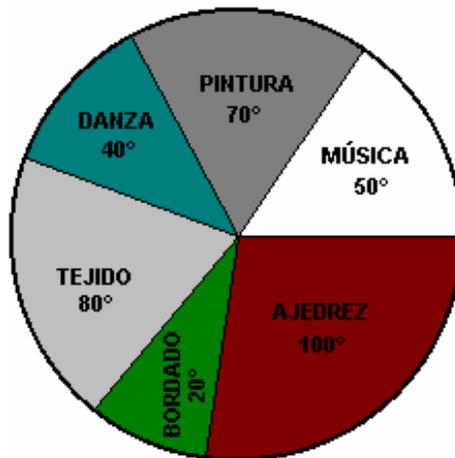
De la misma forma calcula los demás valores.

Los resultados que debiste obtener son los que se muestran en la tabla:

CATEGORÍAS	FRECUENCIA (f _i)	GRADOS
Música	250	50°
Pintura	350	70°
Danza	200	40°
Tejido	400	80°
Bordado	100	20°
Ajedrez	500	100°
Total	1800	360°

Para representarlos en el círculo, medimos los sectores con un transportador y trazamos las divisiones.

1. Alumnos inscritos en los cursos dominicales.



Gráfica No 7

Para hacer resaltar las regiones, éstas se somborean de diferente tono.

Una forma más útil de representar las regiones de cada categoría es distribuyéndolos en un porcentaje y para ello la regla de tres que se plantea, se relaciona 360° representa el 100% y en base a ésta se determina el porcentaje que le corresponde a los grados de cada categoría (curso); veamos el planteamiento:

$$\begin{array}{l} 360^\circ \text{ es a } 100 \\ \text{como } 50^\circ \text{ es a } 100 \end{array} \qquad \begin{array}{l} 360^\circ \text{ es a } 100 \\ \text{como } 70^\circ \text{ es a } x \end{array}$$

$$\begin{array}{ccc} 360^\circ & \xrightarrow{\quad} & 100 \\ & \swarrow \quad \searrow & \\ 50^\circ & & x \end{array}$$

$$\begin{array}{ccc} 360^\circ & \xrightarrow{\quad} & 100 \\ & \swarrow \quad \searrow & \\ 70^\circ & & x \end{array}$$

$$\begin{aligned} 360^\circ x &= 50^\circ (100) \\ x &= \frac{50^\circ (100)}{360^\circ} = \frac{500}{36} \\ x &= 13.9\% \end{aligned}$$

$$\begin{aligned} 360^\circ x &= 70^\circ (100) \\ x &= \frac{70^\circ (100)}{360^\circ} = \frac{700}{36} \\ x &= 19.4\% \end{aligned}$$

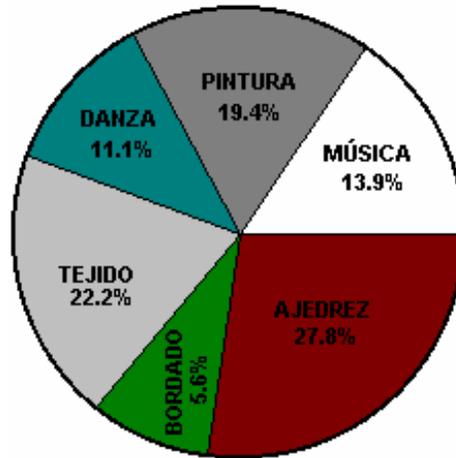
De la misma forma, calcula los demás valores.

Los resultados que debiste obtener son los de la siguiente tabla:

CATEGORÍAS	FRECUENCIA (f _i)	GRADOS	%
Música	250	50°	13.9
Pintura	350	70°	19.4
Danza	200	40°	11.1
Tejido	400	80°	22.2
Bordado	100	20°	5.6
Ajedrez	500	100°	27.8
Total	1800	360°	100.0

Para la representación gráfica se divide el círculo con un transportador de acuerdo a los grados que le corresponde a cada categoría, pero ahora se registra el porcentaje.

1. Alumnos inscritos en los cursos dominicales:



Gráfica No. 8

Información obtenida por el coordinador de los cursos dominicales.

ACTIVIDAD DE REGULACION

Del siguiente problema representa los datos en una gráfica circular indicando el porcentaje correspondiente a cada categoría.

Un transporte cargó su camión con los siguientes productos:

Producto A - 450 kg.

Producto B - 300 kg.

Producto C - 500 kg.

Producto D - 1600 kg.

Producto E - 1750 kg.

Gráficas de Barras

Otra forma de representar gráficamente las puntuaciones de un problema, es mediante una gráfica de barras.

Para construir una gráfica de barras se trazan ejes coordenados; en el eje horizontal se representan los valores de la variable y se traza un segmento perpendicular por cada valor. Si se usan barras, éstas deberán tener el mismo ancho de la base.

En el eje vertical se representa la frecuencia de cada clase usando una escala conveniente para cada puntuación. La frecuencia marca la altura de cada segmento perpendicular o barra.

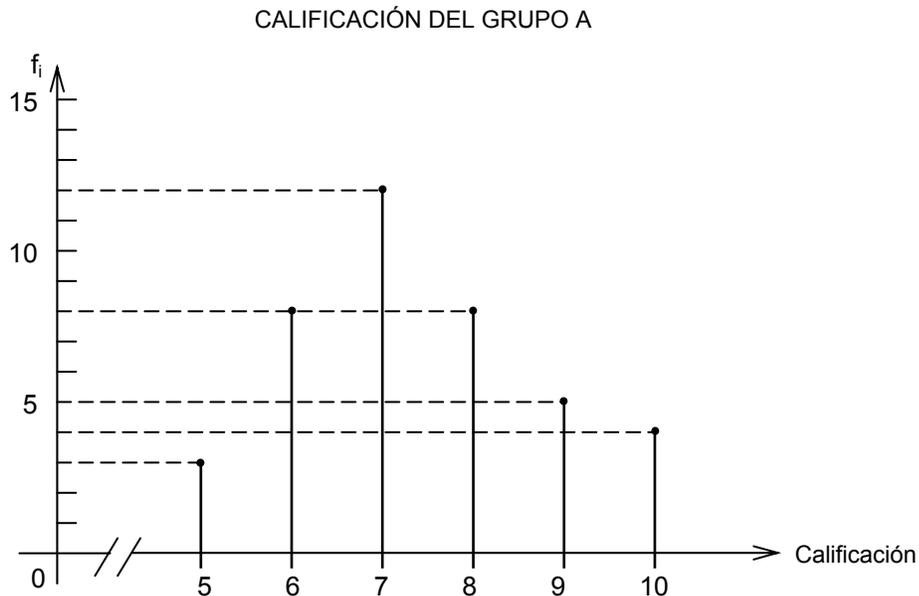
11. Analicemos el siguiente problema:

El profesor de estadística usa una escala de clasificación del 5 al 10, aprobatoria a partir de 6. Cualquier calificación reprobatoria se refleja en un 5.

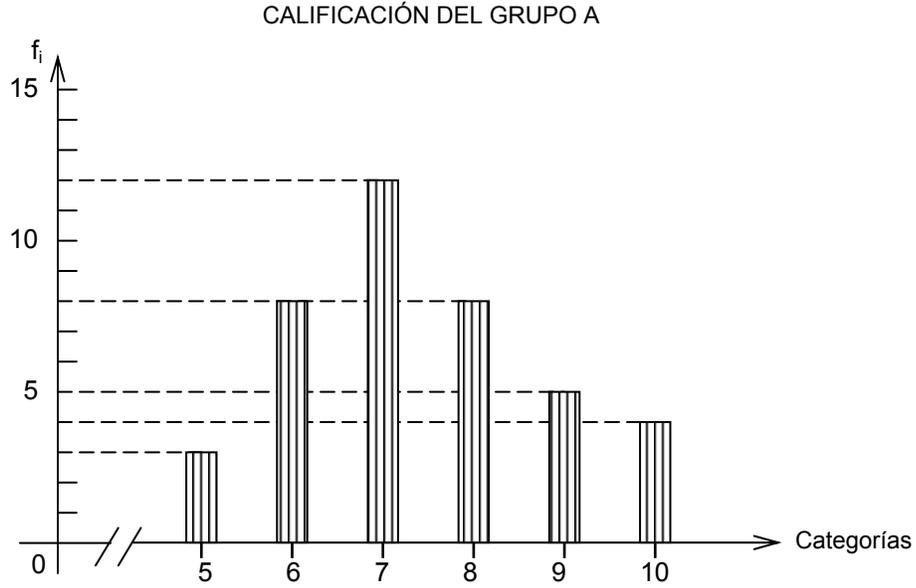
Las calificaciones obtenidas del grupo A son las que se muestran en la tabla siguiente:

Calificación	5	6	7	8	9	10
No. de Alumnos	3	8	12	8	5	4

Representa la información en una gráfica de barras, usando segmentos y barras.



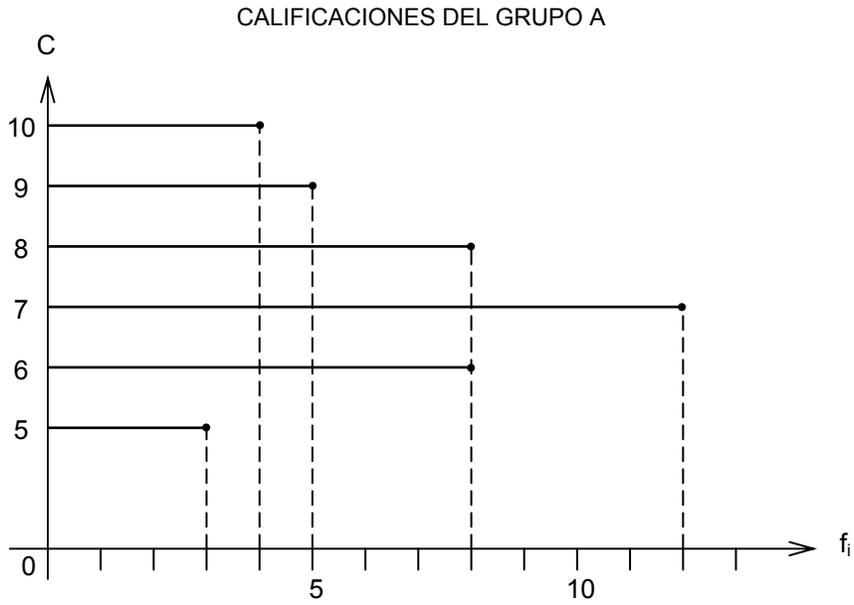
Gráfica No. 9
Diagrama de segmentos verticales



Gráfica No. 10
Diagrama de barras

Otra forma de representar la información es mediante barras horizontales, en este caso se invierten las escalas, es decir, en el eje horizontal se representan las frecuencias y en el vertical los valores de la variable.

Veamos cómo quedaría la gráfica del problema anterior.



Gráfica No. 11

ACTIVIDAD DE REGULACION

Representa el mismo problema en una gráfica de barras horizontales y contesta las siguientes preguntas:

- a) ¿Cuál es la variable del problema? _____
- b) ¿Qué tipo de variable es? _____
- c) ¿En cuántas categorías se clasificó la variable? _____
- d) ¿Cuál es la frecuencia de la más alta calificación? _____

Pictograma

Es otra forma de representación gráfica de la información de un determinado problema.

En este tipo de gráfica se usa la imagen de la variable, por ejemplo:

- a) Árboles: si la variable representa árboles.
- b) Libros: si la variable representa libros.
- c) Casas: si la variable representa casas.
- d) Figuras de personas: si la variable son personas, etcétera.

12. Veamos el siguiente ejemplo:

El regente de la ciudad de México envió a 5 delegaciones el siguiente número de árboles para reforestar su zona:

- | | |
|------------------------|--------------|
| 1. Gustavo A. Madero | 5000 árboles |
| 2. Venustiano Carranza | 3000 árboles |
| 3. Cuauhtemos | 4000 árboles |
| 4. Cuajimalpa | 6000 árboles |
| 5. Benito Juárez | 7000 árboles |

La representación gráfica de la variable mediante un pictograma es la que se muestra a continuación:

DELEGACIÓN	No. DE ÁRBOLES (en miles)
Gustavo A. Madero	
Venustiano Carranza	
Cuauhtemoc	
Cuajimalpa	
Benito Juárez	

Para poder determinar la cantidad de elementos que representa el pictograma, se le asigna un valor a la figura representativa. En nuestro ejemplo cada árbol representa a mil de éstos.

ACTIVIDAD DE REGULACION

En el siguiente problema representa los datos de la variable mediante un pictograma.

Se hizo una investigación del crecimiento poblacional del municipio de Teteles de Ávila Castillo, Estado de Puebla, la información obtenida de los archivos son:

AÑO	No. HABITANTES
1985	8000
1986	9000
1987	10000
1988	11000
1989	12000
1990	13000

Gráfico de Líneas

Una forma de representar gráficamente los valores de la variable de un problema en estudio, es mediante un gráfico de líneas.

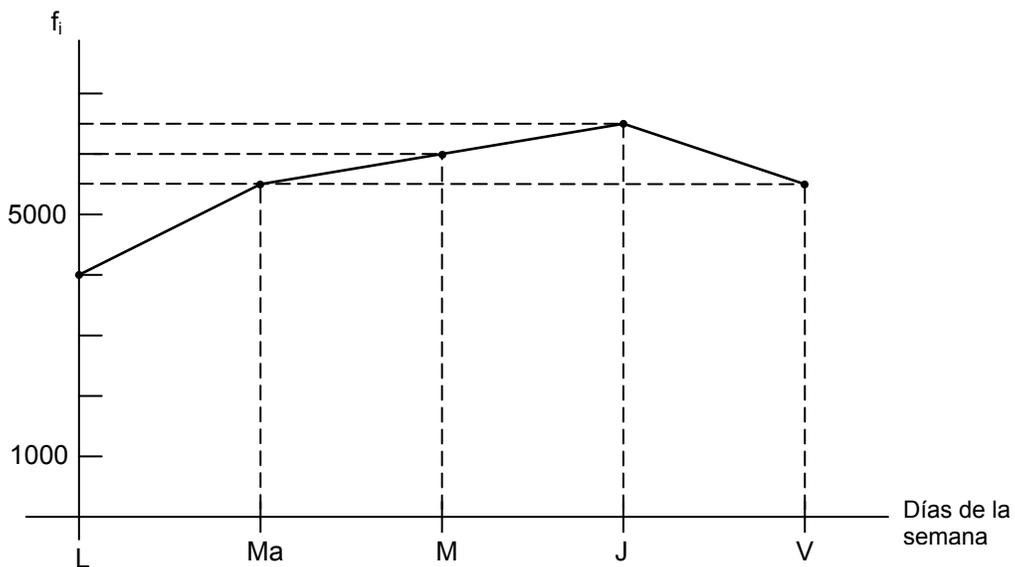
Para trazar la gráfica de líneas se usa el plano coordenado; en el eje horizontal se representa a la variable y en el eje vertical la frecuencia. Se determinan los puntos de corte del valor de la variable con su frecuencia y se unen estos puntos obteniéndose la gráfica de línea que nos muestra con claridad los cambios que experimentó la variable.

14. Veamos el siguiente ejemplo:

La producción en la segunda semana de septiembre obtenida por un obrero calificado que opera un torno revólver, con el que produce tornillos de cuerda milimétrica, es la siguiente:

DÍA	CANTIDAD
Lunes (L)	4000
Martes (Ma)	5500
Miércoles (M)	6000
Jueves (J)	6500
Viernes (V)	5500

Producción de tornillos del obrero Pedro Rodríguez obtenida en la segunda semana de septiembre.



Gráfica No. 12

Elaboró el jefe de taller de la empresa.

- a) ¿Cuál es la variable del problema? _____
- b) ¿Qué tipo de variable es? _____
- c) ¿En cuántas categorías se agrupa la variable? _____
- d) ¿En qué día hubo mayor producción? _____

ACTIVIDAD DE REGULACION

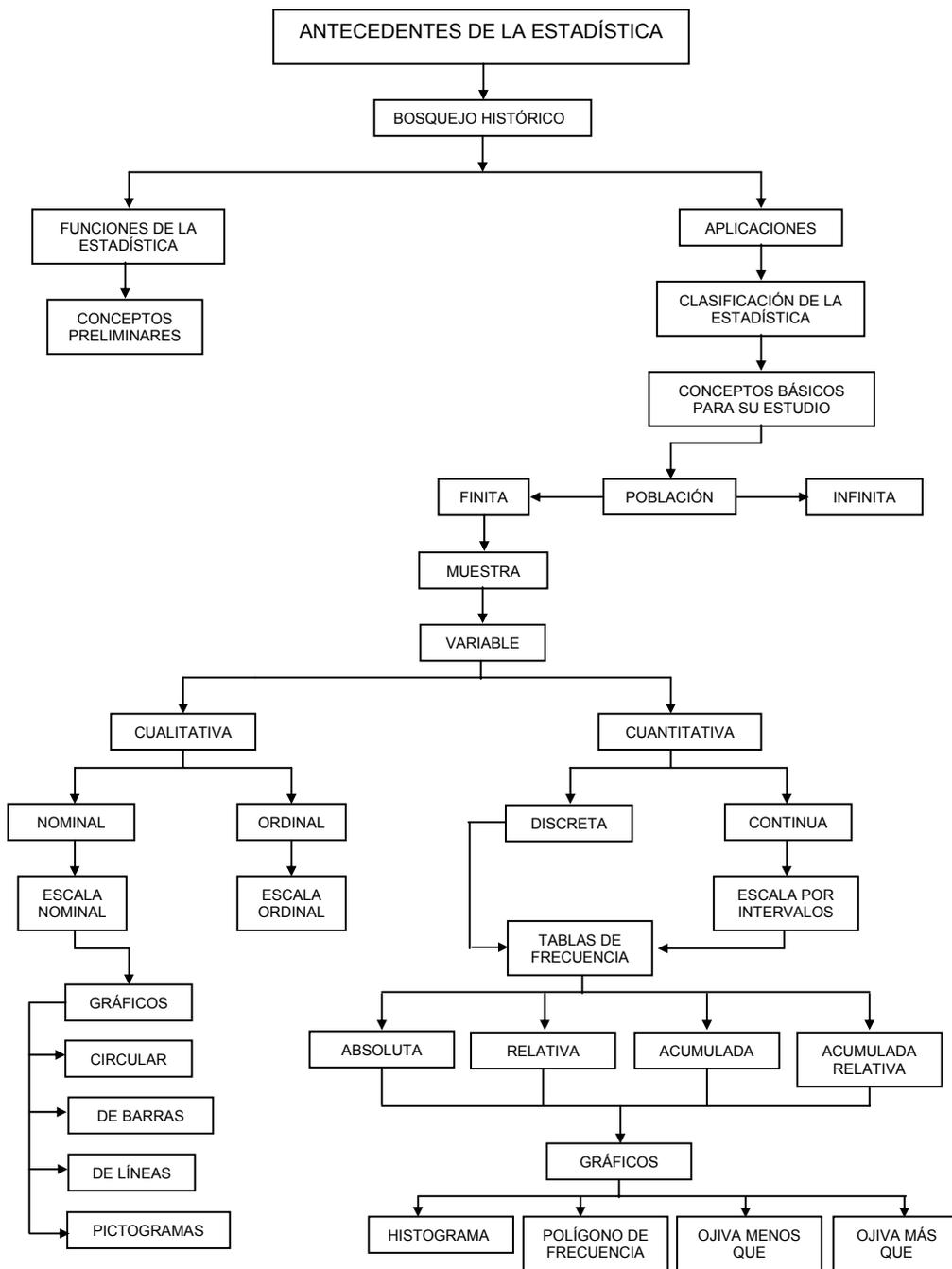
Traza un gráfico de líneas y contesta las preguntas que se plantean.,

El número de adultos alfabetizados en el estado de Hidalgo son los siguientes:

AÑO	No. ADULTOS
1985	40,000
1986	50,000
1987	60,000
1988	55,000
1989	50,000

- a) ¿Cuál es la variable del problema? _____
- b) ¿Qué tipo de variable es? _____
- c) ¿En cuántas categorías se agrupa la variable? _____
- d) ¿En qué año hubieron mejores resultados? _____

RECAPITULACIÓN



ACTIVIDADES DE CONSOLIDACIÓN

Se han preparado estos ejercicios para que reafirmes tus conocimientos adquiridos en este fascículo, resuélvelos y si tienes alguna duda, consulta a tu profesor o asesor académico.

1. Se tiene una urna con 50 esferas numeradas progresivamente. El experimento consiste en sacar una esfera, anotar su número y devolverla a la urna.
 - a) ¿Cuál es la variable? _____
 - b) ¿La variable es cualitativa o cuantitativa? _____
 - c) Si podemos sacar una esfera por tiempo indefinido, entonces ¿cómo es la población? _____
 - d) Si únicamente podemos hacer 100 extracciones, entonces, ¿cómo es la población? _____

2. En el Plantel 2 del Colegio de Bachilleres ingresaron 8000 estudiantes. De estos se seleccionaron 500 para estudiar las características que indican. Escribe en la línea si la variable asociada está definida por una expresión o un número; determina si es cualitativa o cuantitativa e indica el tipo de escala.
 - a) Estado de salud de cada alumno. _____
 - b) Estatura de cada uno. _____
 - c) Número de estudiantes que ingresan. _____
 - d) Sexo de cada estudiante. _____
 - e) ¿Cuál es y cómo es la población? _____
 - f) ¿Cuál es la muestra? _____
 - g) Se escogen los 10 primeros lugares en puntuación. _____

3. En la Fuerza Aérea Mexicana, los aviadores toman un examen médico general en cada 150 horas de vuelo, por lo que se analizan los expedientes de cada piloto para citarlos a su examen. Los resultados obtenidos se muestran en la siguiente tabla correspondientes a las horas de vuelo.

222,	220,	199,	212,	201,	217,	214,	213,	212,	215,	214,
196,	192,	208,	207,	212,	214,	209,	202,	206,	196,	202,
219,	194,	206,	204,	194,	211,	216,	209,	212,	216,	206,
201,	201,	210,	200,	206,	201,	205,	217,	219,	212,	207,
211,	204,	212,	212,	205,	207,	214,	211,	220,	205,	201,
218,	193,	199,	206,	209,	205,	205,	221,	216,	211,	204,
215,	203,	213,	207,	209,	197,	197,	199,	208,	209,	218,
204,	207,	202,	209,	195,	207,	209,	205,	210,	197,	209,
211,	208,	211,	198,	217,	209,	199,	192,	218,	193,	208,
209,	194,	211,								

- Ordena los datos en una tabla de frecuencias.
- Calcula el rango de los datos.
- Calcula el número de clase mediante la aplicación de la fórmula de Sturges $K = 1 + 3.322 \log n$
- Calcula la amplitud del intervalo $A = \frac{R}{K}$
- Determina los límites de clase.
- Determina los límites reales de clase.
- Determina los puntos medios de clase M_i
- Determina la frecuencia f_i
- Determina la frecuencia relativa f_r
- Determina la frecuencia acumulada F_i
- Determina la frecuencia acumulada relativa.
- Traza el histograma y el polígono de frecuencia.
- Traza el polígono de frecuencia acumulada relativa (ojiva).
- Traza el polígono de frecuencia acumulada menos que.
- Traza el polígono de frecuencia acumulada más que.
- Traza un gráfico circular.

4. Los subsidios que el gobierno federal distribuyó en el año de 1993 se muestran en la siguiente tabla.

PARAESTATAL	MILLONES EN NUEVOS PESOS
Azúcar	3,000
Conasupo	4,000
I.M.S.S.	6,000
C.F.E.	8,000
Pemex	10,000
A.S.A.	8,000
D.I.M.	7,000
F.A.M.	6,000
Ecología	4,000
C.R.M.	2,000

Con los datos de la tabla:

- Construye un gráfico de barras verticales.
- Un gráfico de barras horizontales.

5. La venta de libros de la librería de Cristal de los 5 años anteriores fueron los que se muestran en la siguiente tabla.

	1988	1989	1990	1991	1992
Matemáticas	500	1000	1500	2000	3000
Física	1000	500	2500	2000	1500
Química	500	1500	500	500	2000
Filosofía	500	500	1000	500	500
Leyes	1000	1500	2000	1000	500

Con los datos de la tabla elabora un pictograma.

6. El consumo en miles de litros de agua de una industria durante los primeros 10 meses del año de 1993 son los que se muestran en la siguiente tabla.

MES	No. DE LITROS EN MILES
Enero	50
Febrero	58
Marzo	62
Abril	70
Mayo	80
Junio	85
Julio	80
Agosto	70
Septiembre	65
octubre	60

Representa los datos de la tabla en un diagrama lineal.

AUTOEVALUACIÓN

Los resultados que debiste obtener en la solución de los problemas anteriores son los siguientes:

Si algunos resultados no coinciden, entonces regresa a la unidad correspondiente, repásala y aclara tus dudas con tu profesor o tu consultor académico.

1.
 - a) La variable es el número de la esfera.
 - b) Es cuantitativa.
 - c) La población es infinita.
 - d) La población es finita.

2.
 - a) Se define por una expresión (estado de salud) es de tipo cualitativo, es nominal, es una variable discreta y la escala con que se cuantifica es nominal.
 - b) Se define por un número (número de centímetros), es de tipo cuantitativo, es una variable continua y su escala con que se cuantifica es por intervalos.
 - c) Se define por un número, es de tipo cualitativo, es una variable discreta y su escala de medición es por intervalos.
 - d) Se define por una expresión (sexo) es de tipo cualitativo nominal, es discreta: su escala de medición es nominal.
 - e) La población está constituida por 8000 estudiantes, por lo tanto es una población finita.
 - f) La muestra está constituida por 500 alumnos.
 - g) La variable está definida por una expresión (lugar que ocupó), es de tipo cualitativo ordinal, la escala con que se cuantifica es ordinal y es una variable discreta.

3. DATOS ORDENADOS

a)

192	192	193	193	194	194	194	195	196	196
197	197	197	198	199	199	199	199	200	200
201	201	201	201	201	202	202	202	203	204
204	204	204	205	205	205	205	205	206	206
206	206	206	207	207	207	207	207	207	208
208	208	208	209	209	209	209	209	209	209
209	210	210	211	211	211	211	211	211	211
212	212	212	212	212	212	212	213	213	213
214	214	214	214	215	215	215	216	216	216
217	217	218	218	218	219	219	220	221	222

b) $R = X_n - X_1 = 222 - 192 = 30$

R = Rango

c) $K = 1 + 3.322 \log(102) = 1 + 3.322(2.008) = 1 + 6.67 = 7.67 \approx 8$

K = 8

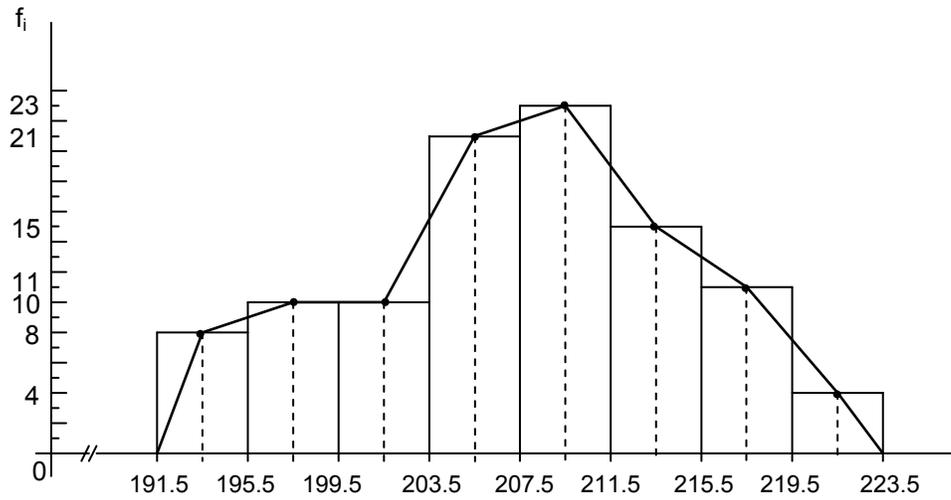
K = Número de clases

d) $A = \frac{R}{K} = \frac{30}{8} = 3.75 \approx 4$

A = Amplitud de clase

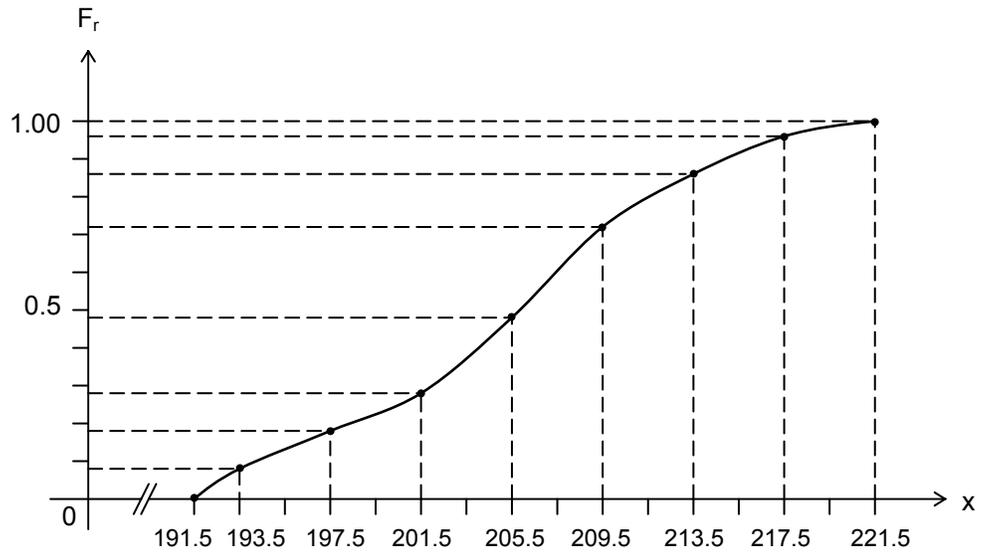
e)	f)	g)	h)	i)	j)	k)
Límite de clase	Límite real de clase	Punto Medio de clase	Frecuencia Absoluta f_r	Frecuencia Relativa f_r	Frecuencia Acumulada F_i	Frecuencia Relativa Acumul. F_r
192-195	191.5-195.5	193.5	8	0.078	8	0.078
196-199	195.5-199.5	197.5	10	0.098	18	0.176
202-203	199.5-203.5	201.5	10	0.098	28	0.274
204-207	203.5-207.5	205.5	21	0.21	49	0.484
208-211	207.5-211.5	209.5	23	0.23	72	0.714
212-215	211.5-215.5	213.5	15	0.15	87	0.864
216-219	215.5-219.5	217.5	11	0.10	98	0.964
220-223	219.5-223.5	221.5	4	0.036	102	1.00
Total			102	1.00		

l) HISTOGRAMA Y POLÍGONO DE FRECUENCIA



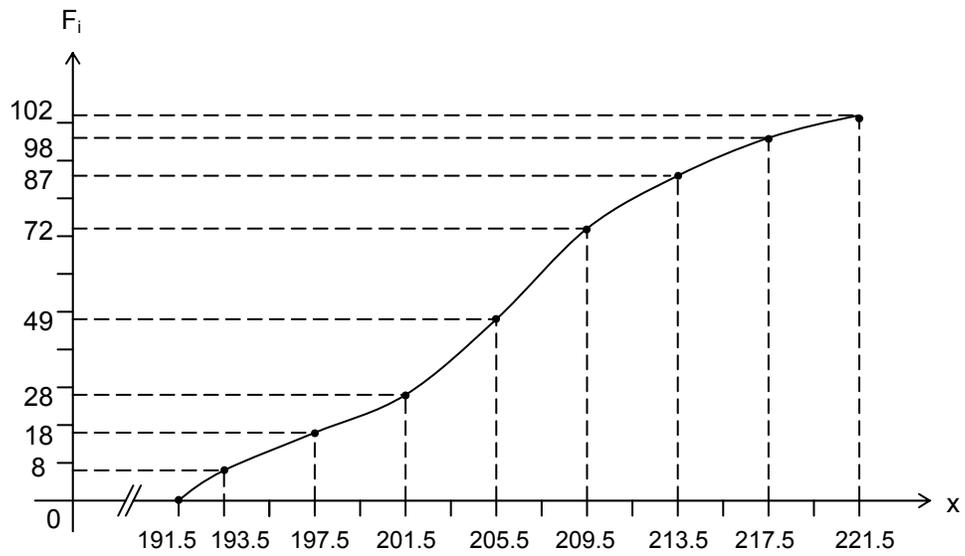
m) POLÍGONO DE FRECUENCIA ACUMULADA RELATIVA

PUNTO MEDIO DE CLASE	FRECUENCIA ABSOLUTA ACUMULADA F_i	FRECUENCIA RELATIVA ACUMULADA F_R
191.5	0	0.0
193.5	8	0.078
197.5	18	0.176
201.5	28	0.274
205.5	49	0.484
209.5	72	0.714
213.5	87	0.864
217.5	98	0.964
221.5	102	1.00



n) OJIVA MENOS QUE

PUNTO MEDIO DE CLASE	FRECUENCIA ABSOLUTA f_i	FRECUENCIA ABSOLUTA ACUMULADA F_i
191.5	0	0
193.5	8	8
197.5	10	18
201.5	10	28
205.5	21	49
209.5	23	72
213.5	15	87
217.5	11	98
221.5	4	102



menos que → Horas de vuelo

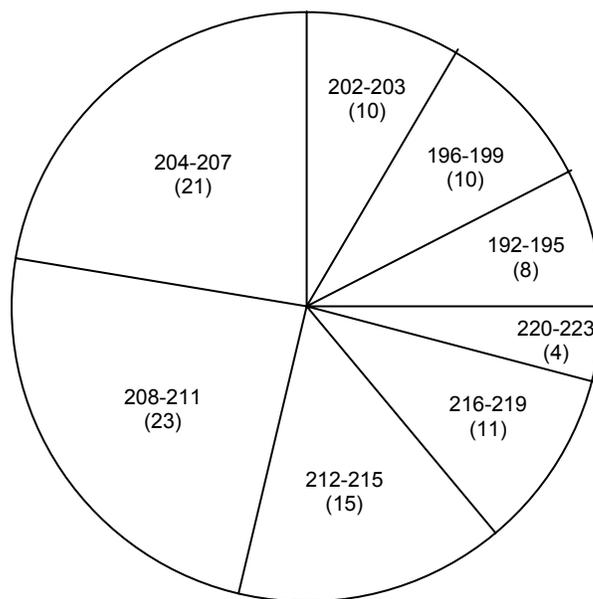
o) OJIVA MÁS QUE

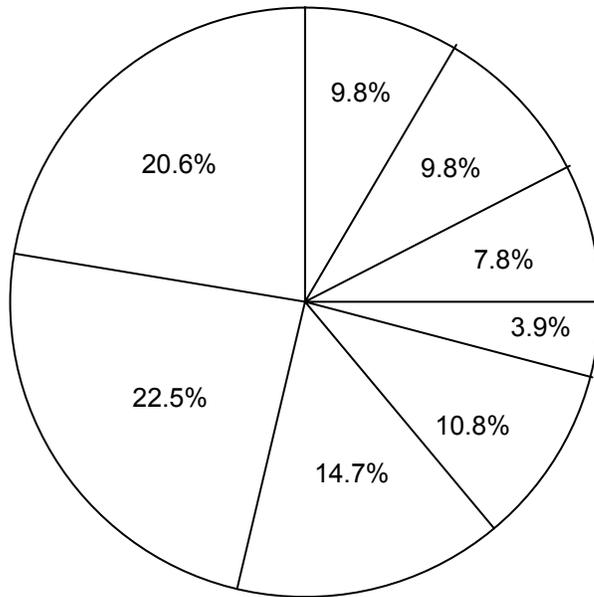
PUNTO MEDIO DE CLASE	FRECUENCIA ABSOLUTA f_i	FRECUENCIA ACUMULADA MÁS QUE
193.5	8	102
197.5	10	94
201.5	10	84
205.5	21	74
209.5	23	53
213.5	15	30
217.5	11	15
221.5	4	4
223.5	0	0



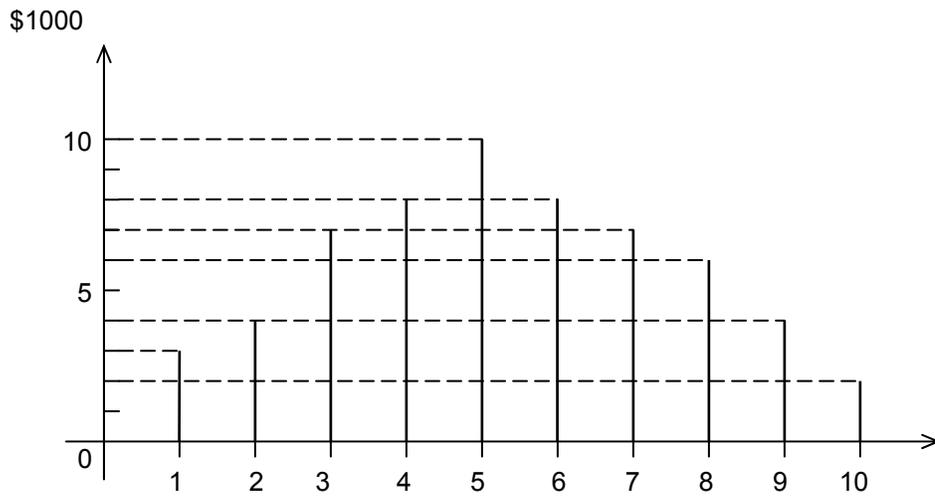
más que → Horas de vuelo

p) DIAGRAMA CIRCULAR DE LAS HORAS DE VUELO

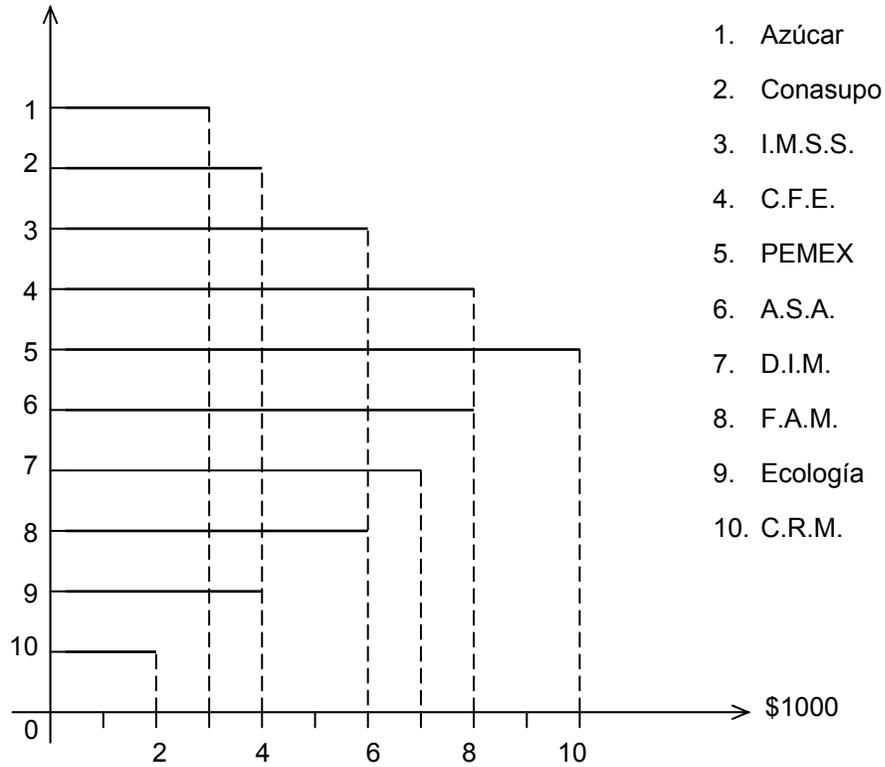




4. a) GRÁFICO DE BARRAS VERTICALES



b) GRÁFICO DE BARRAS HORIZONTAL



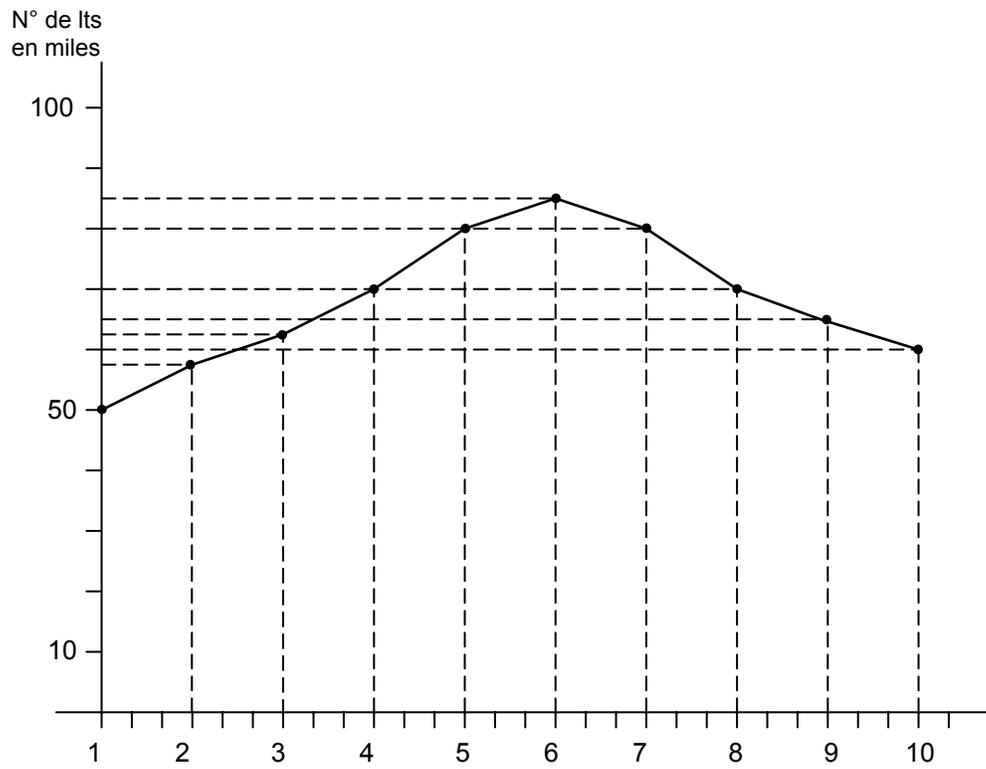
5) PICTOGRAMA

Cada libro representa 500 libros



	1988	1989	1990
Matemáticas	1 icon	2 icons	3 icons
Física	2 icons	1 icon	5 icons
Química	1 icon	3 icons	1 icon
Filosofía	1 icon	1 icon	2 icons
Leyes	2 icons	3 icons	4 icons

6) DIAGRAMA LINEAL



- | | |
|------------|---------------|
| 1. Enero | 6. Junio |
| 2. Febrero | 7. Julio |
| 3. Marzo | 8. Agosto |
| 4. Abril | 9. Septiembre |
| 5. Mayo | 10. Octubre |

BIBLIOGRAFÍA CONSULTADA

FREUD E., John, Williams J., Frank, Perles M., Benjamin. Estadística para la administración. Editorial Prentice Hall Hispanoamericana, S.A. México

G., HOEL Paul. Estadística elemental. Editorial CECSA. México, 1976.

KREYSZING, Erwin. Introducción a la estadística matemática. Editorial Limusa. México, 1981.

LINCOLN L., Chao. Introducción a la estadística. Editorial CECSA,. México, 1985

MENDENHALL, William. Introducción a la probabilidad y la estadística. Traductor Segami Carlos.

PARSEN, Emanuel. Teoría moderna de probabilidad y sus aplicaciones. Editorial Limusa. México, 1973.

SEYMOUR Lipschutz. Ph. D. Teoría y problemas de probabilidad. Editorial McGraw-Hill. México, 1973.

STEVENSON, William J. Estadística para administración y economía. Editorial Harla. México, 1981.

TARO, Yamane. Estadística. Editorial Harla, 3ª. edición. México, 1979.

WILLOUGHBY, Stephens. Probabilidad y estadística. Publicaciones Culturales, S.A. México, 1983.

