



Secretaría de Educación Pública
Subsecretaría de Educación Media Superior
Dirección General del Bachillerato



Quinto semestre

Probabilidad y estadística I



Probabilidad y estadística I

Quinto semestre



La Patria (1962),
Jorge González Camarena.

Esta obra ilustró la portada de los primeros libros de texto. Hoy la reproducimos aquí para mostrarte lo que entonces era una aspiración: que estos libros estuvieran entre los legados que la Patria deja a sus hijos.

Estimada, estimado estudiante del Telebachillerato Comunitario, este libro fue elaborado pensando en ti, forma parte de una colección que incluye todas las asignaturas del plan y los programas de estudio.

En su elaboración participaron profesionales y especialistas en distintas disciplinas, quienes tomaron en cuenta tus necesidades e inquietudes. En estos libros hallarás contenidos y actividades que contribuirán a que logres un mejor desempeño ahora que cursas la Educación Media Superior.

Tenemos la certeza de que con los materiales didácticos del Telebachillerato Comunitario, con el apoyo de tus maestras, maestros y con tu propio esfuerzo, tendrás un mejor aprovechamiento escolar y contribuirás al bienestar de tu comunidad y de México.

Te deseamos éxito en esta importante etapa de tu formación.

DISTRIBUCIÓN GRATUITA, PROHIBIDA SU VENTA

Probabilidad y estadística I

SEP

SECRETARÍA DE
EDUCACIÓN PÚBLICA



Telebachillerato comunitario. Quinto semestre.

Probabilidad y Estadística I

Secretaría de Educación Pública

Emilio Chuayffet Chemor

Subsecretaría de Educación Media Superior

Rodolfo Tuirán Gutiérrez

Dirección General del Bachillerato

Carlos Santos Ancira

Autoras:

Martha Rocío Estrada Rivera

Xochitl Itzel Hernández

Asesoría académica:

Cayetano Quincunx Avelleyra

Servicios editoriales:

Facultad Latinoamericana de Ciencias Sociales (FLACSO)

Coordinación General: Lorenzo Gómez Morin Fuentes

Editores Responsables: José Ángel Quintanilla D'Acosta y Mónica Lobatón Díaz

Diseño y diagramación, material fotográfico e iconografía:

Instituto de Gestión e Innovación Educativa (IGIE)

José Pedro Cortés Xiqui, Maritza Sosa Amenyro y Vanessa Alejandra Valadez Gutiérrez

Primera edición

Derechos Reservados. Secretaría de Educación Pública, 2015 ©

Argentina 28, Centro, 06020, México, D.F.

ISBN: 978-607-8229-81-9

Impreso en México

Prefacio

Estimado estudiante, el libro que tienes en tus manos fue elaborado pensando en ti, en tus necesidades e inquietudes, como un instrumento que te apoye ahora que estudias el bachillerato. En sus páginas encontrarás contenidos y actividades que son fundamentales para que paso a paso, puedas alcanzar las metas que la asignatura te propone para este semestre.

A ti te toca, ahora, sacarle el mayor provecho a este libro, que es fruto del esfuerzo de un grupo de profesores y especialistas. Si lo haces tu amigo, lo aprovechas al máximo y lo combinas con el apoyo de tu maestro y de los demás recursos didácticos que están a tu alcance, seguramente ampliarás tus competencias y habilidades para construir un mejor futuro para ti, y contribuir al desarrollo de tu comunidad, de tu estado y de nuestro México.

Te deseamos el mayor de los éxitos en esta importante etapa de tu formación, el bachillerato.

Tabla de contenido

Probabilidad y Estadística I

Presentación general	8
¿Cómo está estructurado este libro?	10
Simbología	13
¿Con qué conocimientos cuento?	14

Bloque I. Comprendes y describes la variabilidad estadística y sus aplicaciones

Conceptos básicos	22
La Estadística	22
Población y muestra	24
Estadística descriptiva e inferencia estadística	27
Variables y variabilidad	28
Clasificación de las variables	33
Recolección de datos	33

Bloque II. Describes y representas datos de forma tabular y gráfica

Organización de los datos	40
¿Para qué te sirve tener organizada la información?	40
Formas de organización, dependiendo del tipo de datos	41
Tipos de frecuencias: absoluta, relativa, acumulada y relativa acumulada	43
Análisis de datos a través de tablas	46
Partes de una tabla	54
Tipos de tablas y su uso	61
Análisis de datos a través de gráficas	63
Histograma de frecuencia de clases	63
Polígono de frecuencias	64

Bloque III. Aplicas la Estadística descriptiva

Estadígrafos de tendencia central	75
La media aritmética	75
La mediana	76
Cuartiles, deciles, percentiles	77
La moda	79
Relaciones empíricas entre la media, la mediana y la moda	80
Estadígrafos de dispersión	89
Rangos	90
Desviaciones	92

Bloque IV. Analizas la teoría de conjuntos y sus aplicaciones

Relaciones y operaciones entre conjuntos	114
Teoría de probabilidad	122
Glosario.....	129
Apéndice	130
Referencias	155

Presentación general

La asignatura Probabilidad y estadística I es parte de tu formación propedéutica de bachillerato y pertenece al campo disciplinar de Matemáticas. Conforme al marco curricular común, tiene la finalidad de propiciar el desarrollo de la creatividad y el pensamiento lógico y crítico, mediante procesos de razonamiento, argumentación y construcción de ideas, que promueven el desarrollo de distintas competencias para la resolución de problemas que trasciendan el ámbito escolar.

Probabilidad y estadística I se ubica en el quinto semestre del plan de estudios del bachillerato general que ha establecido la Secretaría de Educación Pública (SEP); tiene relación con las asignaturas: Matemáticas I, II, III y IV, así como con Taller de Desarrollo Comunitario.

Te invitamos a aprovechar al máximo este libro, el cual está integrado por una serie de contenidos y actividades de aprendizaje, a través de los cuales desarrollarás conocimientos, habilidades, actitudes y valores para crecer como persona y como ciudadano, capaz de resolver y comprender situaciones de la vida cotidiana a través del lenguaje científico y matemático, así como a construir nuevos conocimientos y compartirlos con quienes te rodean.

Enfoque para el desarrollo de competencias



¿Qué es una competencia?

La educación basada en competencias pretende que adquieras los conocimientos, desarrolles las habilidades y muestres las actitudes necesarias para resolver y enfrentarte a situaciones de la vida diaria de mejor manera.

En el telebachillerato comunitario, se consideran tres tipos de competencias: genéricas, disciplinares y profesionales. En ésta asignatura Probabilidad y Estadística I se promueve el desarrollo de las dos primeras que podrás observar; las competencias genéricas que se abordan se mencionan a continuación, y al inicio de cada bloque las disciplinares.

Competencias genéricas	Atributos
1. Te conoces y te valoras a ti mismo y abordas problemas y retos teniendo en cuenta los objetivos que persigues.	<ul style="list-style-type: none">Analizas críticamente los factores que influyen en la toma de decisiones.
4. Escuchas, interpretas y emites mensajes pertinentes en distintos contextos mediante la utilización de medios, códigos y herramientas apropiados.	<ul style="list-style-type: none">Expresas ideas y conceptos mediante representaciones lingüísticas, matemáticas o gráficas.
5. Desarrollas innovaciones y propones soluciones a problemas a partir de métodos establecidos.	<ul style="list-style-type: none">Sigues instrucciones y procedimientos de manera reflexiva, comprendiendo cómo cada uno de tus pasos contribuyen a alcanzar un objetivo.Ordenas información de acuerdo con categorías, jerarquías y relaciones.Construyes hipótesis y diseñas y aplicas modelos para probar su validez.Sintetizas evidencias obtenidas mediante la experimentación para producir conclusiones y formular nuevas preguntas.
6. Sustentas una postura personal sobre temas de interés y relevancia general, considerando otros puntos de vista de manera crítica y reflexiva.	<ul style="list-style-type: none">Eliges las fuentes de información más relevantes para un propósito específico y discriminas entre ellas de acuerdo con su relevancia y confiabilidad.
7. Aprendes por iniciativa e interés propio a lo largo de la vida.	<ul style="list-style-type: none">Articulas saberes de diversos campos y estableces relaciones entre ellos y tu vida cotidiana.
8. Participas y colaboras de manera efectiva en equipos diversos.	<ul style="list-style-type: none">Asumes una actitud constructiva, congruente con los conocimientos y habilidades con los que cuentas, en distintos equipos de trabajo.

¿Cómo está estructurado este libro?



Inicio de cada bloque

Cada bloque comienza con un esquema en el que se muestran los objetos de aprendizaje, los productos y las competencias disciplinares que se abordarán.

Posteriormente se presenta una breve introducción en donde se indica de qué trata y cómo vas a trabajar.

Asimismo, se presenta el propósito del bloque, es decir, las metas y los desempeños que debes lograr.

Para identificar qué tanto sabes del tema y cuáles son las áreas por mejorar, se propone una evaluación diagnóstica, que además te permitirá conocer tu nivel en las competencias a desarrollar durante el curso.

Bloque I Comprende y describes la variabilidad estadística y sus aplicaciones

Bloque I

10 HORAS

Objetos de aprendizaje que se abordan

1. Conceptos básicos.
 - 1.1 La estadística.
 - 1.2 Población y muestra.
 - 1.3 Estadística descriptiva e inferencia estadística.

Desempeños esperados al concluir el bloque

- Valora a la estadística como una herramienta matemática que le permite tomar decisiones para organizar, resumir datos y transmitir resultados de forma significativa.
- Distingue las ramas de la estadística para identificar su aplicación en diferentes situaciones.
- Reconoce las características de una población y las técnicas de recolección de datos para aplicarlas en situaciones hipotéticas.
- Valora las ventajas que tiene el emplear las diversas técnicas de muestreo para el análisis de los datos de una población o muestra.
- Comprende, identifica y describes las variables como atributos de interés de los datos provenientes de una población o muestra para reconocer su comportamiento.

Competencias disciplinares del campo de las matemáticas:

- Estructura ideas y argumentos de manera clara, coherente y sintética, relacionadas con la estadística descriptiva e inferencial.
- Argumenta el uso de la estadística descriptiva e inferencial.
- Identifica las ideas clave en un texto sobre las técnicas de recolección de datos e infiere conclusiones a partir de ellas.
- Analiza las relaciones entre dos o más variables de un proceso social o natural para determinar o estimar su comportamiento.
- Expresa ideas y conceptos sobre las relaciones entre los datos recolectados de una población para determinar o estimar su comportamiento.
- Analiza críticamente los factores que influyen en su toma de decisiones en la organización de datos.
- Ordena los datos de una población de acuerdo a categorías, jerarquías y relaciones.
- Analiza las relaciones entre las variables y los datos en un proceso social o natural para determinar o estimar el comportamiento de la población objetivo.

Productos de aprendizaje

- Actividad 1 y 2. Selección de tema y como abordarlo de la mejor manera.
- Actividad 3. Condiciones para la validez de la información.
- Actividad 4. Listado de población y muestra.
- Actividad 5. Análisis de tabla.

Comprende y describes la variabilidad estadística y sus aplicaciones

Introducción

En este bloque vas a aprender qué es la estadística, cuáles son sus partes y las aplicaciones que puede tener en tu vida cotidiana.

En primer lugar, aprenderás que son una variable y una constante, qué es la variabilidad y qué tipos de variables existen.

En segundo lugar, manejarás algunos de los instrumentos que te permiten recolectar datos y organizarlos para que posteriormente puedas interpretarlos. Por último podrás representar la información recabada en tablas o gráficas y hacer una primera interpretación de los datos.

Para que tú aprendizaje sea significativo, a lo largo de todo el bloque te presentaremos información relacionada con tu vida diaria.

Al término del bloque obtendrás como producto la representación gráfica de la información recogida, así como una primera interpretación de dicha información.

Mapa de objetos de aprendizaje

```
graph TD; Estadística --> Técnicas[Técnicas de recolección de datos]; Estadística --> Variables; Técnicas --> Descriptiva; Técnicas --> Inferencial; Variables --> Descriptiva; Variables --> Inferencial;
```



Desarrollo del bloque

Esta parte es fundamental, pues aquí encontrarás el contenido general y disciplinar que necesitas para acercarte al tema de la Estadística.

A lo largo del bloque se intercalan estrategias didácticas de aprendizaje, actividades acompañadas de imágenes, ejemplos, preguntas detonadoras y evaluaciones. Todo está relacionado con los contenidos y las competencias a desarrollar. También encontrarás algunos apoyos de estudio como cápsulas con datos interesantes y cuadros al margen del texto para reforzar tu aprendizaje; por ejemplo:

Aplicas la estadística descriptiva

Actividad de aprendizaje 5

A partir de lo que acabamos de comentar, calcula las frecuencias absoluta, absoluta acumulada, relativa y relativa acumulada de tu cuadro.

¡Sorpresa!

Cuando hayas concluido, léete a la vista los dos datos que necesitas para calcular la media aritmética:

$$\text{media aritmética} = \frac{\text{total de la columna 3}}{\text{total de la columna 1}}$$

Actividad de aprendizaje 6

Calcula la media aritmética de las edades de las personas entrevistadas de tu comunidad, organizadas por clases y frecuencias.

Sabías que...

La media aritmética de un conjunto de datos organizados en clases y frecuencias se obtiene:

1. Multiplicando las observaciones de cada clase por su marca de clase respectiva.
2. Sumando los resultados obtenidos para cada una de las clases de su clasificación de frecuencias.
3. Dividiendo el resultado entre el total de las observaciones de todas las clases.

Si la fórmula para calcular la media aritmética de datos no organizados es:

$$\bar{x} = \frac{\sum xij}{n}$$

1. Glosario, definiciones y términos para apoyar la comprensión del texto.

2. Imágenes, que te ayudarán a una mejor comprensión de los conceptos.

3. Figuras, que te permitirán realizar las actividades de aprendizaje.

4. Datos interesantes, que facilitan la relación de los contenidos con tu vida diaria.

Bloque III Aplicas la estadística descriptiva

Sabías que...

- La media es un estadígrafo de tendencia central que mide magnitud.
- La mediana es un estadígrafo de tendencia central que mide posición.
- La moda es un estadígrafo de tendencia central que mide la frecuencia.

Aprende más

Sabiendo todo esto, vamos ahora a aprender a calcular la moda cuando los datos están organizados en clases y frecuencias.

Lo primero que tienes que hacer es ubicar en qué clase está la moda. ¡Eso es muy fácil!

En tu cuadro, en la columna de observaciones es la que tiene la mayor cantidad y ahí está.

Ahora sabes que la fórmula dice:

$$M^o = Ei + \left(\frac{A_1}{A_1 + A_2} \right) C$$

Donde:

- L_i = es el límite inferior de la clase en la que cae la moda (mismo concepto para calcular la mediana).
- A_1 = en cuántos es mayor el número de observaciones de la clase modal, respecto a la de la clase anterior.
- A_2 = en cuántos es mayor el número de observaciones de la clase modal, respecto a la de la clase siguiente.
- C = tamaño del intervalo de una clase modal (mismo concepto que para calcular la media).

¿Cómo está estructurado este libro?



Cierre de bloque

Al terminar cada tema se te pedirá una actividad y un producto final para que puedas evaluar cuánto has avanzado y qué áreas de oportunidad tienes; asimismo se te pedirá analizar, investigar, reflexionar y argumentar acerca de los temas señalados.

El libro incluye actividades de aprendizaje para que evalúes tu desempeño en el logro de las competencias. Al finalizar cada actividad puedes consultar la retroalimentación de la misma en el apartado *Apéndice* al final del libro. Ten presente que cada actividad debe concretarse en una evidencia que irás recopilando en tu cuaderno y concentrando para la evaluación del curso.

Bloque II Describes y representas datos, tablas y gráficos

Cierre de bloque II

Reflexiona sobre lo aprendido

La recopilación de la información te ayuda a analizar un problema, a conocer una población.

En este bloque estudiamos cómo partiendo de una masa de datos construimos un cuadro de conteo que nos ayuda a organizar la información para posteriormente resumirla en clases y construir la tabla de distribución de frecuencias.

Además aprendiste que hay cuatro tipos de frecuencias, las frecuencias absolutas, las frecuencias relativas y las frecuencias acumuladas, tanto absolutas como relativas.

Acuéstate que las frecuencias relativas te permiten comparar dos poblaciones de tamaños diferentes y que las gráficas valen más que mil números pues tienen un impacto visual que te permite sacar conclusiones rápidas de la masa de datos.

Por último, aprendiste a construir un histograma y un polígono de frecuencias y a diferenciarlas. Analizaste cómo el área de una distribución de frecuencias muestra características de una masa de datos que no se observan a simple vista.

Bloque I Reconoces lugares geométricos

Autoevaluación

Lee detenidamente las preguntas y responde colocando una (X) en el nivel de avance que consideras que has logrado a lo largo del bloque I.

Interpretación del nivel de avance:

100-90% = Lo logré de manera independiente
80-70% = Requerí apoyo para construir el aprendizaje
60-50% = Fue difícil el proceso de aprendizaje y sólo lo logré parcialmente
40% o menos = No logré el aprendizaje

Contenidos	Nivel de avance			
	100-90%	80-70%	60-50%	40% o menos
Identificas para qué te sirve la estadística en tu vida diaria.				
Comprendes la diferencia entre estadística descriptiva e inferencial.				
Identificas los diferentes instrumentos de recolección de datos.				
Reconoces en los textos diferentes tipos de variables.				

Contenidos	Nivel de avance			
	100-90%	80-70%	60-50%	40% o menos
Analizas críticamente la información que se te presenta.				
Distingues los elementos de una población y una muestra.				
Construyes ejemplos de población y muestra.				
Explicas con tus propias palabras la diferencia entre variable discreta y continua, y variable dependiente e independiente.				

Los contenidos y las actividades se presentan de una manera atractiva. Aprovecha cada pregunta, contenido y actividades, ya que incidirán en tu crecimiento personal, familiar y social.

Trabaja con tu profesor y con tus compañeros; acércate a ellos, resuelvan dudas y aprendan juntos; date la oportunidad de construir con ellos este viaje. Esperamos que el curso te sea interesante y fructífero.



Simbología que facilitará tu proceso de aprendizaje

Diseño Instruccional



Aprende más



Actividad de aprendizaje

Apoyos para reforzar el aprendizaje



Glosario



Reflexionemos sobre la actividad



Sabías que...

¿Con qué conocimientos cuento?

Para que se te facilite la comprensión de los contenidos que estamos por mostrarte, te proponemos hacer un breve repaso de lo que aprendiste en Matemáticas II sobre estadística y probabilidad, en Matemáticas III sobre la utilización de distintas formas de la ecuación de una recta, en Matemáticas IV sobre ecuaciones lineales simples, así como elementos de aritmética básica, como la regla de tres y los porcentajes. A continuación te presentamos algunos ejercicios de diagnóstico con el propósito de que repases un poco antes de iniciar.

Esta evaluación tiene 7 reactivos, si aciertas entre 7 y 6 tu nivel es excelente y puedes iniciar el estudio de la asignatura. Si obtienes entre 5 y 4 tu nivel es muy bueno, aunque sería recomendable que repasaras tus clases de Matemáticas. Si obtienes 3 o menos, te proponemos hacer un repaso general de los temas que se tratan en esta evaluación ¿estás listo?

¡Iniciemos!

Recuerda que siempre existen diferentes procedimientos para solucionar un problema. Tú decides cuál es la mejor forma para tí.

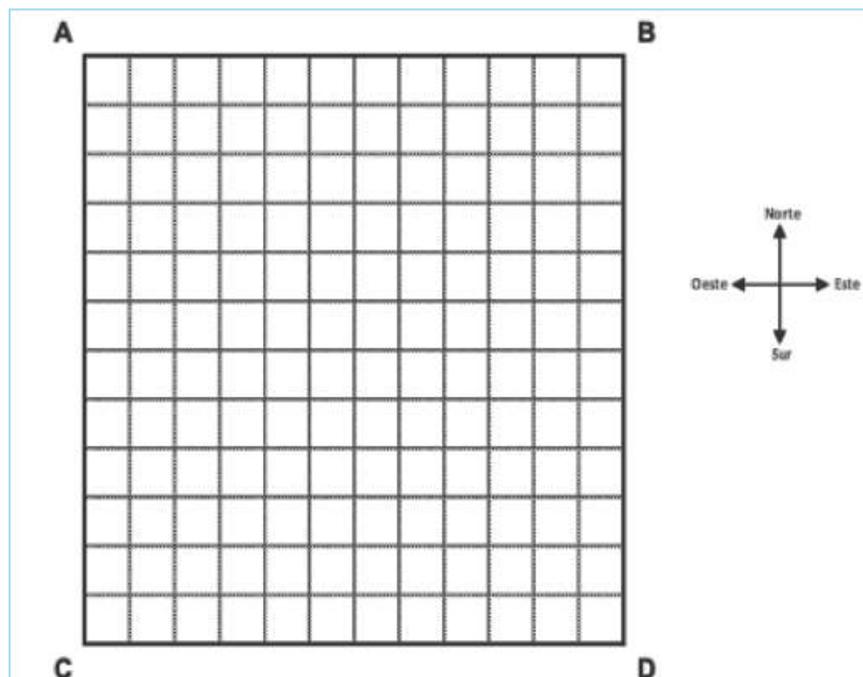
1. Un medicamento para un niño se debe administrar a razón de dos gotas por cada 10 kg de peso, si el niño pesa 22.8 kg, ¿cuántas gotas deberá administrarle su mamá?

2. La presidencia municipal ha regalado al Telebachillerato 12 botes de pintura azul, cada uno de $\frac{1}{2}$ L. Con ellos, entre tus compañeros y tú pintaron la barda que tiene 90 metros de largo y 80 centímetros de altura. Calcula cuántos botes de 2 litros de pintura azul serán necesarios para pintar todo el edificio que mide 15 m de altura y 40 m de longitud.

3. Veinte de los 30 compañeros del salón se fueron de viaje. ¿Qué proporción de alumnos salieron de viaje?

4. Para el festejo de la fiesta patronal de la localidad, un alumno del telebachillerato comunitario compró y donó dos artículos con el fin de venderlos y recaudar fondos. Si se vendieron a 75 pesos cada uno de ellos, con una ganancia de 10%, ¿cuánto le costaron ambos artículos?

5. La tía Rosario quiere reforestar su terreno y te pidió que calcularas cuántos árboles necesita para cubrirlo completamente ubicando cada árbol a 10 metros de distancia del otro. Te dibujó un plano con las características de su terreno. En cada extremo hay un poste (A, B, C, D) que sostiene la cerca a lo largo del perímetro del terreno, por lo que no se puede sembrar árboles en el perímetro. Cada cuadro mide 10 x 10 metros.



- Localiza y dibuja en el plano la casa de la tía Rosario que está a 20 m hacia al Este del poste A y a 40 m del poste B hacia el oeste. La casa es un cuadrado que tiene un área de 360 m^2 .
- ¿Cuántos árboles necesitaste para reforestar el terreno de la tía Rosario, sin considerar el área ocupada por la casa?

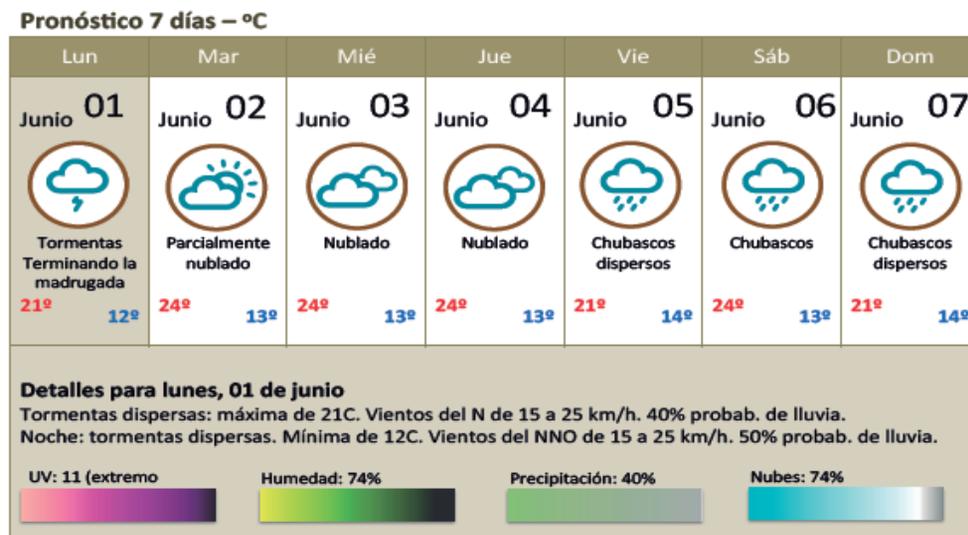
¿Con qué conocimientos cuento?

6. Se tienen un par de termómetros, uno graduado en grados fahrenheit ($^{\circ}\text{F}$) y el otro en grados centígrados ($^{\circ}\text{C}$). Para averiguar cómo están relacionadas estas dos escalas se hacen las mediciones de temperatura mostradas en la siguiente tabla:

$^{\circ}\text{F}$	50	131	203
$^{\circ}\text{C}$	10	55	95

Identifique la ecuación algebraica de la relación entre ambos termómetros:

- $^{\circ}\text{C} - ^{\circ}\text{F} = 0$
 - $^{\circ}\text{C} - ^{\circ}\text{F} + 32 = 0$
 - $9/5 ^{\circ}\text{C} - ^{\circ}\text{F} + 32 = 0$
 - $9/5 ^{\circ}\text{F} - ^{\circ}\text{C} + = 0$
7. A continuación se presentan las condiciones del clima en una semana:



Calcula la probabilidad de que la siguiente semana se repitan las mismas condiciones climatológicas en:

- Días soleados
- Días mayormente soleados
- Días parcialmente nublados
- Días lluviosos

¿Cómo viste esta evaluación? Cuando la hayas terminado, puedes consultar las propuestas de solución en la sección al final del libro llamada Apéndice. Si ya lo hiciste, estás listo para empezar con el primer bloque de la asignatura. ¡Que te diviertas!

BLOQUE I

Comprendes y describes la variabilidad estadística y sus aplicaciones



¿Qué aprenderás y cómo organizarás tu estudio?

Bloque I

10
HORAS

Objetos de aprendizaje que se abordan

Conceptos básicos.
La Estadística
Población y muestra
Estadística descriptiva e inferencia estadística
Variables y variabilidad
Recolección de datos

Desempeños esperados al concluir el bloque

- Valoras a la Estadística como una herramienta matemática que te permite tomar decisiones para organizar, resumir datos y transmitir resultados de forma significativa.
- Distingues las ramas de la Estadística para identificar su aplicación en diferentes situaciones.
- Reconoces las características de una población y las técnicas de recolección de datos para aplicarlas en situaciones hipotéticas.
- Valoras las ventajas que tiene el emplear las diversas técnicas de muestreo para el análisis de los datos de una muestra o una población.
- Comprendes, identificas y describes las variables como atributos de interés de los datos provenientes de una población o muestra para reconocer su comportamiento y diferencias.

Competencias disciplinares del campo de las matemáticas

- Estructuras ideas y argumentos de manera clara, coherente y sintética, relacionadas con la Estadística descriptiva e inferencial.
- Argumentas el uso de la Estadística descriptiva e inferencial.
- Identificas las ideas clave en un texto sobre las técnicas de recolección de datos e infieres conclusiones a partir de ellas.
- Analizas las relaciones entre dos o más variables de un proceso social o natural para determinar o estimar su comportamiento.
- Expresas ideas y conceptos sobre las relaciones entre los datos recolectados de una población para determinar o estimar su comportamiento.
- Analizas críticamente los factores que influyen en la organización de datos.
- Ordenas los datos de una población de acuerdo con categorías, jerarquías y relaciones.
- Analizas las relaciones entre las variables y los datos, en un proceso social o natural para determinar o estimar el comportamiento de la población objetivo.

Productos de aprendizaje

- Durante este bloque realizarás un primer documento de interpretación de la información presentada, que pondrá de manifiesto el desarrollo de tus competencias.

Introducción

En este bloque vas a aprender qué es la Estadística, cuáles son sus partes y las aplicaciones que puede tener en tu vida cotidiana.

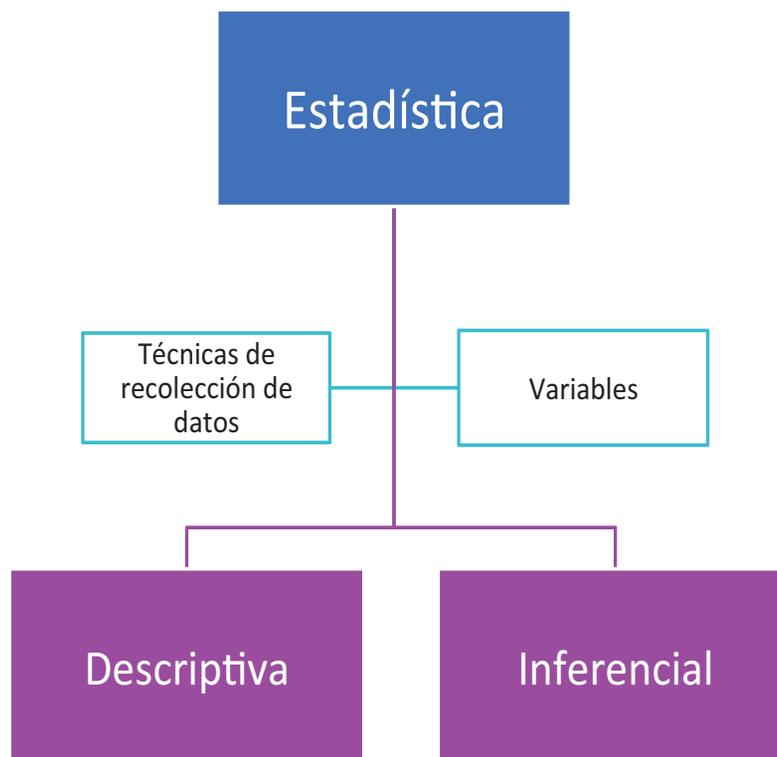
En primer lugar, aprenderás qué son una variable y una constante, qué es la variabilidad y qué tipos de variables existen.

En segundo lugar, manejarás algunos de los instrumentos que te permiten recolectar datos y organizarlos para que posteriormente puedas interpretarlos. Por último, podrás representar la información recabada en tablas o gráficas y hacer una primera interpretación de los datos.

Para que tu aprendizaje sea significativo, a lo largo de todo el bloque te presentaremos información relacionada con tu vida diaria.

Al término del bloque obtendrás como producto la representación gráfica de la información recogida, así como una primera interpretación de dicha información.

Mapa de objetos de aprendizaje





Para iniciar, reflexiona

Al leer el periódico o escuchar las noticias encuentras información como la siguiente:

México, el tercer país con más jóvenes que no estudian ni trabajan: OCDE.



(Notimex) - México tiene el tercer porcentaje más alto de jóvenes que no estudian ni trabajan, entre los 34 países de la Organización para la Cooperación y Desarrollo Económicos (OCDE), sólo por debajo de Turquía e Israel.

El informe Panorama de la Educación 2013 de la OCDE revela que el 24.7% de los jóvenes mexicanos de 15 a 29 años no tienen trabajo ni asisten a la escuela, un nivel que se ha mantenido por casi una década. Del porcentaje de jóvenes que no estudian ni tienen empleo en México, las mujeres encabezan la cifra nacional, según la organización.

Disponible: <http://mexico.cnn.com/nacional/2013/06/25/mexico-el-tercer-pais-con-mas-jovenes-que-no-estudian-ni-trabaja-ocde> Consultado el 21 de abril de 2015.

De 100 partidos, Chivas sólo ha ganado 26

En el último centenar de duelos el 'Rebaño' acumula casi la mitad de derrotas lo que explica la triste actualidad del conjunto tapatío.

MÉXICO, D.F.- Los números que Chivas presenta en los últimos cien partidos disputados en la liga local son un claro reflejo del porqué hoy se ubican en la parte baja de la tabla porcentual. En el último centenar de sus cotejos, Chivas ha acumulado 44 derrotas y 30 empates, registrando la alarmante cifra de sólo 26 triunfos, de los cuales 14 fueron en el Omniflife, por 12 en campo ajeno.

Su estadio, a pesar de ser el escenario en el que más victorias han conseguido, también representa un cheque en blanco para aquellas escuadras que lo visitan, pues en 35 ocasiones sus rivales lograron cosechar al menos un punto del campo del Ominilife. Chivas ha regalado 78 unidades repartidas en derrotas y empates jugando como local.

Disponible en: SIPSE en línea. "De 100 partidos, Chivas sólo ha ganado 26", Por Héctor Ortiz de Milenio. 19 de agosto de 2014 en <http://sipse.com/deportes/liga-mx-crisis-chivas-de-guadalajara-porcentaje-de-descenso-107823.html> Consultado el 17 de abril de 2015.

MXN/ USD - Pesos mexicanos por dólar

Fecha	Apertura	Cierre	Dif.%	Máximo	Mínimo	Volumen
28/04/2015	15,3664	15,2681	-0,65	15,4000	15,1873	15.300.417
27/04/2015	15,3910	15,3673	-0,15	15,4339	15,2903	15.399.976
26/04/2015	15,4113	15,3905	0,11	15,4255	15,3597	15.424.017
25/04/2015	15,3735	15,3735	0,00	15,3735	15,3735	154.072
24/04/2015	15,3410	15,3735	0,20	15,4440	15,2955	15.407.067

Disponible en: <http://www.finanzas.com/divisas/usd-mxn/> Consultado el 21 de abril de 2015.

Toda esta información son datos estadísticos y es parte de nuestra vida cotidiana, de lo que platicamos con nuestras familias, amigos, compañeros de la escuela o del trabajo. Contar con ella y saber analizarla nos ayuda a tomar decisiones. Por ejemplo, si te preguntan si el equipo ganará su próximo partido, ¿cuál sería tu respuesta en términos de probabilidad que tiene ese equipo de ganar su próximo partido?



Sabías que...

La Estadística ha estado presente en la historia de la humanidad desde las antiguas civilizaciones. Por ejemplo, en el Antiguo Egipto, debido a las inundaciones del río Nilo, se efectuaban trabajos censales que permitían conocer el reparto de la propiedad y de los bienes, para que fueran restituidos después de las inundaciones. También los griegos levantaban censos de población y de propiedad. Por su parte, en la época del Imperio romano se realizaban censos de bienes y de personas de los pueblos sometidos con el objeto de recolectar impuestos (Bonilla, 1995).

En México el primer registro sobre personas, pueblos y terrenos conquistados es el conocido con el nombre de Suma de Visitas de Pueblos por Orden Alfabético, conservado en la Biblioteca Nacional de Madrid, producido a mediados del siglo XVI por frailes de diversas órdenes; este documento constituye un catastro de las propiedades de los indígenas, una nómina de tributos y un padrón de habitantes de cada una de las 907 jurisdicciones políticas que conformaban el total de las tierras conquistadas y colonizadas.



Censo: el estudio de todos los elementos de una población. En México se realiza el censo de población cada 10 años.



Aprende más

Conceptos básicos

La Estadística

La **Estadística** es una ciencia formal. Esto quiere decir que, como las matemáticas y la lógica, construye y comprueba nociones abstractas. (Recuerda la clasificación de las ciencias que viste en la asignatura Introducción a las Ciencias Sociales).

La Estadística te ayuda a recopilar, analizar, interpretar y presentar información muy diversa: desde los resultados de los equipos de fútbol y las calificaciones de los estudiantes del telebachillerato comunitario, hasta el número de días que ha llovido en tu comunidad durante los últimos cinco años o el número de casos de enfermedades gastrointestinales registrados durante los meses de lluvia.

La información estadística la podemos obtener de diferentes fuentes. Hay fuentes directas, como las entrevistas y encuestas, y fuentes indirectas, como los datos de los periódicos o los informes de organismos nacionales como el Instituto Nacional de Estadística y Geografía (INEGI), el Consejo Nacional de Población (Conapo) e internacionales como la Organización para la Cooperación y Desarrollo Económicos (OCDE) o la Organización Mundial de la Salud (OMS), entre muchas otras.



Estadística: es la ciencia que te ayuda a recopilar, analizar e interpretar información muy diversa. Busca darle sentido a ésta información para ayudar a tomar decisiones.



Actividad de aprendizaje 1

1. Anota en los siguientes renglones cinco temas sobre los cuales te interesaría reunir información detallada, para complementar el diagnóstico que realizaste sobre tu comunidad durante los dos semestres pasados en el marco del Taller de Desarrollo Comunitario.

2. Anota también de qué fuente crees que podrías obtener esta información.

3. Anota, por último, cómo complementarías esta información el diagnóstico con que ya cuentas.

La Estadística busca darle sentido a la información que existe. A partir de la información que recopilas, organizas e interpretas, también te ayuda a tomar decisiones importantes.



Dos ejemplos:

1. Si en una comunidad es elevado el índice de desnutrición, se podría diseñar como apoyo un proyecto de cocina comunitaria.
2. Si se incrementó el número de casos de enfermedades gastrointestinales durante los meses de lluvia, las autoridades sanitarias deberán impulsar medidas como promover campañas para que la población hierva el agua, lave perfectamente los alimentos y aumentar uso de sueros.

Pero lo más relevante de la Estadística es que hoy en día es uno de los soportes más importantes de la investigación y el desarrollo de las ciencias.

Es con base en la estadística que:

- Los científicos estiman los cambios climáticos.
- Los médicos estiman la probabilidad de cura de un enfermo.
- Los politólogos pueden anticipar, por ejemplo, la intención de voto de los jóvenes antes de una elección.



Actividad de aprendizaje 2

Escribe en los siguientes renglones las decisiones que podrías tomar para mejorar en los cinco temas que anotaste en la página anterior y sobre los que podrías reunir información.

1. _____
2. _____
3. _____
4. _____
5. _____

Compara tus respuestas con las de dos de tus compañeros y, en equipo, establezcan cómo esta información reforzará el diagnóstico con el que ya cuentan.

Guarda tus resultados en tu portafolio de evidencias.

Población y muestra

Para hacer una primera aproximación a estos conceptos, supón que dentro de una semana el gobernador de tu estado tendrá una junta con funcionarios de educación federal para tratar el tema de las necesidades de los estudiantes del telebachillerato. Por esta razón, pidió a su equipo de trabajo información sobre los planteles. Sin embargo, los recursos con los que cuenta su equipo sólo alcanzan para encuestar a dos de los 10 telebachilleratos que hay en el estado.



Actividad de aprendizaje 3

¿Crees que la información obtenida en esos dos telebachilleratos pueda ser válida para los demás de tu estado? Argumenta.

¿Bajo qué condiciones esta información podría ser válida para todos los telebachilleratos comunitarios de tu estado? Argumenta.

¿Crees que, incluso, pueda ser válida para los demás telebachilleratos del país? Argumenta.

Reúnete con tus compañeros y discute tus respuestas.

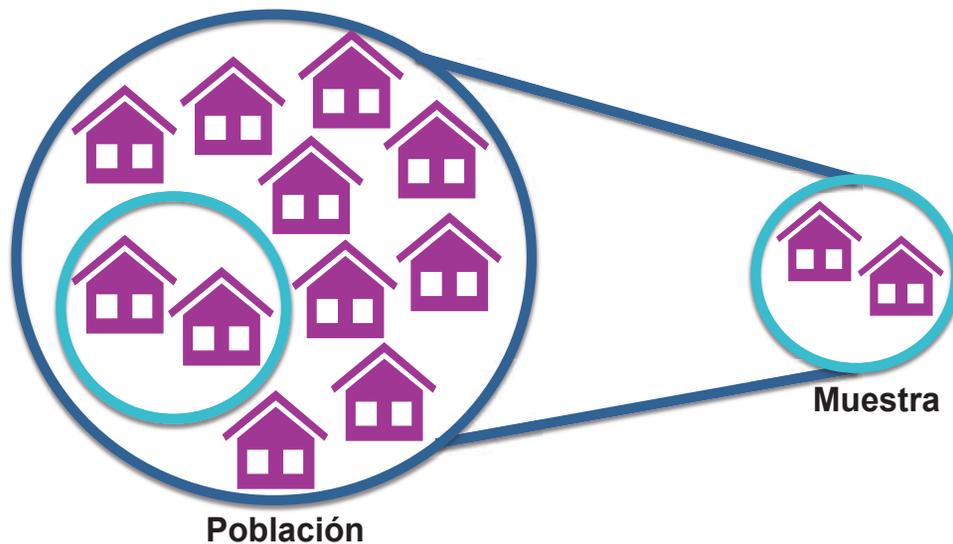
Con la realización de estos ejercicios pudiste identificar la diferencia entre población y muestra. Ahora, con tus compañeros definan otro ejemplo de los conceptos de muestra y población.

Guarda tus resultados en tu portafolio de evidencias.

En Estadística hay dos conceptos fundamentales: la población y la muestra.

La **población** es el conjunto de elementos que te interesa analizar. En nuestro ejemplo las condiciones de todos los telebachilleratos de tu estado.

La **muestra** es un grupo de elementos de esa población. En nuestro ejemplo los dos telebachilleratos encuestados.





Sabías que...

Tras realizar una muestra estadística, los epidemiólogos Richard Doll y Bradford Hill demostraron que fumar era el principal factor de riesgo para desarrollar cáncer pulmonar.



Actividad de aprendizaje 4

Reúnete en equipo y realiza un listado de cinco poblaciones y sus **muestras** que observes en tu entorno cotidiano. Justifica tu respuesta y comparte las conclusiones con tu grupo. Guarda en tu portafolio de evidencias el listado de las poblaciones y sus muestras porque más adelante te será de utilidad.



Muestra: subconjunto representativo de la población a partir del cual se pretende realizar inferencias respecto a la población de donde procede.

Población	Muestra

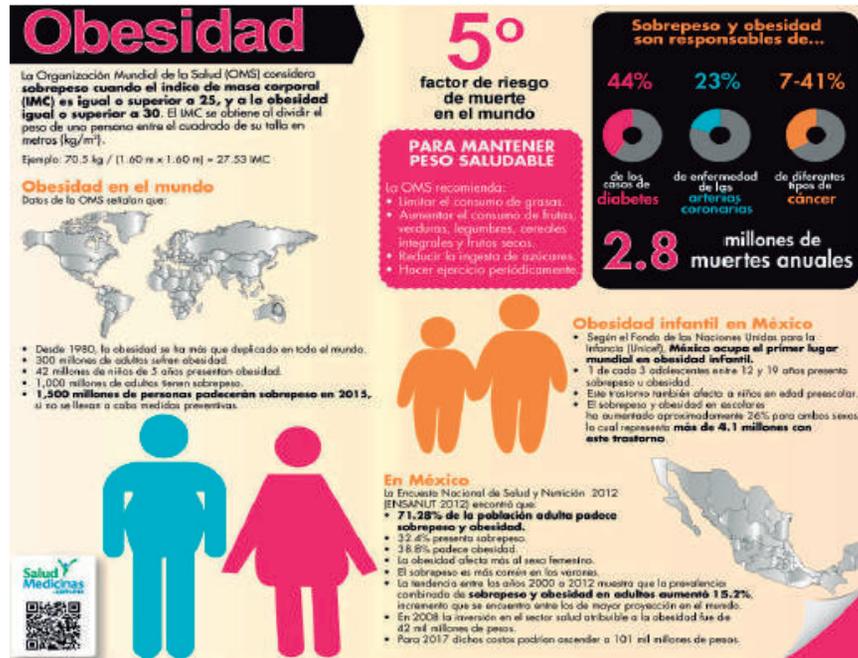


Sabías que...

Del análisis exhaustivo de una muestra puedes inferir el comportamiento de toda la población sin conocerla.

Estadística descriptiva e inferencia estadística

Analiza detenidamente la siguiente información sobre obesidad.



Disponible en: <http://www.saludymedicinas.com.mx/> Consultado el 21 de abril de 2015.



Actividad de aprendizaje 5

Esta actividad busca que profundices en el tema de las fuentes de información recordando lo que trabajaste en el Taller de Desarrollo Comunitario, comenta las preguntas:

1. ¿Cómo crees que la OMS obtuvo estos datos? Discute con tus compañeros.

Es difícil hacer un examen de salud completo a todos los habitantes del planeta, sin embargo, la OMS asegura que: “Desde 1980, la obesidad se ha más que duplicado en todo el mundo.”

2. ¿Cómo puede hacer esa aseveración habiendo en el mundo más de 7 mil millones de habitantes? ¿Nos encuestó a todos?

3. ¿Cuántas personas crees que sea necesario encuestar para hacer afirmaciones como la anterior?

Casi siempre, las poblaciones que queremos analizar son tan grandes que no alcanza ni el tiempo ni los recursos para medir a cada uno de sus integrantes.

Pero además, ¡no es necesario hacerlo!

Si se toman buenas muestras, existe una muy elevada probabilidad de que sus resultados reflejen con gran exactitud las características de toda la **población**. Y fíjate bien que dijimos buenas muestras, no muestras muy grandes.



Por ejemplo, la Encuesta Nacional de Salud y Nutrición 2012 (ENSANUT, 2012) visitó y obtuvo información de 50 528 hogares a nivel nacional distribuidos en las 32 entidades federativas del país, que representan a los 29 429 252 hogares estimados en México para 2012, según las proyecciones de población a partir del Censo 2010. Concluyó que: 71.3% de la población adulta en México padece sobrepeso y obesidad.

Otro ejemplo es la conclusión del Fondo de las Naciones Unidas para la Infancia (UNICEF), en la que México ocupa el primer lugar mundial en obesidad infantil. Para llegar a estas conclusiones, estos organismos utilizaron métodos para obtener muestras confiables y representativas de la población mexicana.

Pues bien, la **estadística descriptiva** se dedica a recopilar, organizar, analizar e interpretar una o varias muestras de la población. A partir de este análisis y siguiendo ciertas reglas, la **inferencia estadística** estudia si las conclusiones de la muestra son válidas para toda la población.

En este curso, trabajarás la parte de Estadística descriptiva y el siguiente semestre verás cómo, el análisis de la muestra, sirve para sacar conclusiones sobre la población.



Estadística descriptiva: conjunto de procedimientos para organizar, resumir y analizar un conjunto de datos.

Inferencia estadística: conjunto de procedimientos cuya finalidad es obtener conclusiones respecto a la población a partir de datos observados en muestras.

Población: conjunto de datos o elementos que interesa analizar.

Variabilidad: se refiere a los cambios que presenta una variable.



Aprende más

Variables y variabilidad

Para iniciar este tema a continuación te presentamos el informe que envió al gobernador del estado su equipo de trabajo respecto a la población de alumnos de telebachillerato:

Comprendes y describes la variabilidad estadística y sus aplicaciones

Tabla 1.1

Telebachillerato comunitario de "Loma Bonita"									
Alumnos	Número de alumnos	Edad promedio	Horas promedio de estudio	Promedio de calificaciones	Número de alumnos unidos (casados, unión libre)	Número de alumnos que son padres o madres	Número de estudiantes que trabajan	Estatura promedio de los varones	Estatura promedio de las mujeres
Primer semestre	18	14	1 ½	10	3	0	1	1.68	1.50
Segundo semestre	14	15	½	7	5	0	1	1.70	1.50
Tercer semestre	24	15	2	7	5	1	2	1.70	1.53
Cuarto semestre	32	17	1 ¾	7	2	1	1	1.65	1.55
Quinto semestre	16	17	1	9	-	1	4	1.72	1.60
Sexto semestre	21	18	3	7	5	2	2	1.72	1.60

Tabla 1.2

Telebachillerato comunitario de "El Platanar"									
Alumnos	Número de alumnos	Edad promedio	Horas promedio de estudio	Promedio de calificaciones	Número de alumnos unidos (casados, unión libre)	Número de alumnos que son padres o madres	Número de estudiantes que trabajan	Estatura promedio de los varones	Estatura promedio de las mujeres
Primer semestre	14	13	0	9	0	0	2	1.60	1.45
Segundo semestre	20	14	0	7	0	1	5	1.61	1.44
Tercer semestre	20	15	2	8	1	1	1	1.65	1.45
Cuarto semestre	35	17	2	6	1	1	2	1.68	1.47
Quinto semestre	29	17	1	6	2	1	2	1.68	1.50
Sexto semestre	35	18	2	7	3	1	2	1.68	1.50



Actividad de aprendizaje 6

1. ¿Cuáles variables de las tablas son importantes para medir el desempeño académico de los alumnos? Justifica tu respuesta.

2. ¿El promedio de calificaciones de los alumnos depende de su estatura? Justifica tu respuesta.

3. ¿Qué puedes observar del comportamiento de las variables a lo largo de los seis semestres? Elige tres y explica su comportamiento.

4. Discute en equipo ¿qué actividades podrían realizar para mejorar el promedio de calificaciones en ambas comunidades?

Con la realización de estos ejercicios podrás identificar los valores que puede tomar una **variable**.

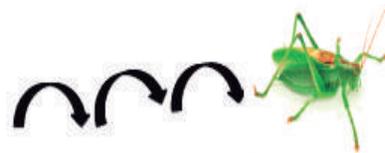


Variable: característica de la población o de la muestra, cuya medida puede cambiar de valor.

Al analizar los valores de estas tablas puedes darte cuenta que no todas las variables manejan el mismo tipo de números.

En términos generales, podemos decir que hay dos tipos de variables: *discretas* y *continuas*.

Una variable discreta es similar a un saltamontes, avanza a saltos y deja sus huellas en unos puntos determinados.



Una variable continua es como un caracol que va dejando un rastro continuo tras de sí.



Las variables discretas son las que se obtienen de contar (número de televisores en los hogares, número de casas en tu localidad, número de niños que asisten a tercer año de primaria, etc.).

Las variables continuas son las que resultan de medir (estatura, peso, talla, distancia de un lugar a otro, etc.).



En las primeras se observan saltos entre un valor y otro, por lo que el cambio se presenta en valores enteros (10 mujeres embarazadas, 5 casos de diabetes). En las segundas no; una variable continua puede tomar cualquier valor (1.69 metros de estatura; 1,604.5 kilómetros de distancia).



Actividad de aprendizaje 7

1. Identifica en las tablas anteriores las variables discretas y continuas.
2. De la siguiente lista de variables identifica las que no son discretas:
 - a) Temperaturas registradas cada hora por el observatorio _____
 - b) Número de hijos de 50 familias _____
 - c) Edades de los mexicanos en el censo _____
 - d) Estatura de los estudiantes de tu escuela _____
 - e) Días del año _____
 - f) Goles en un partido de futbol _____

Estos ejercicios te permiten identificar variables discretas y continuas.



Sabías que...

Cuando el valor de la variable es siempre el mismo, se dice que es una *constante*. Hay variables cuyo valor no se expresa numéricamente. A este tipo de variables se les denomina *variables cualitativas*.

¿Recuerdas que en Matemáticas IV revisaste qué es una variable dependiente y cuál la independiente? Considera dos columnas de la tabla 1.1 y menciona cuál sería la variable dependiente y cuál la independiente. Justifica tu respuesta.



Actividad de aprendizaje 8

1. ¿Crees que exista alguna variable que influya más en el promedio de calificaciones de los alumnos de la primera tabla que en el de los de la segunda? Justifica tu respuesta.

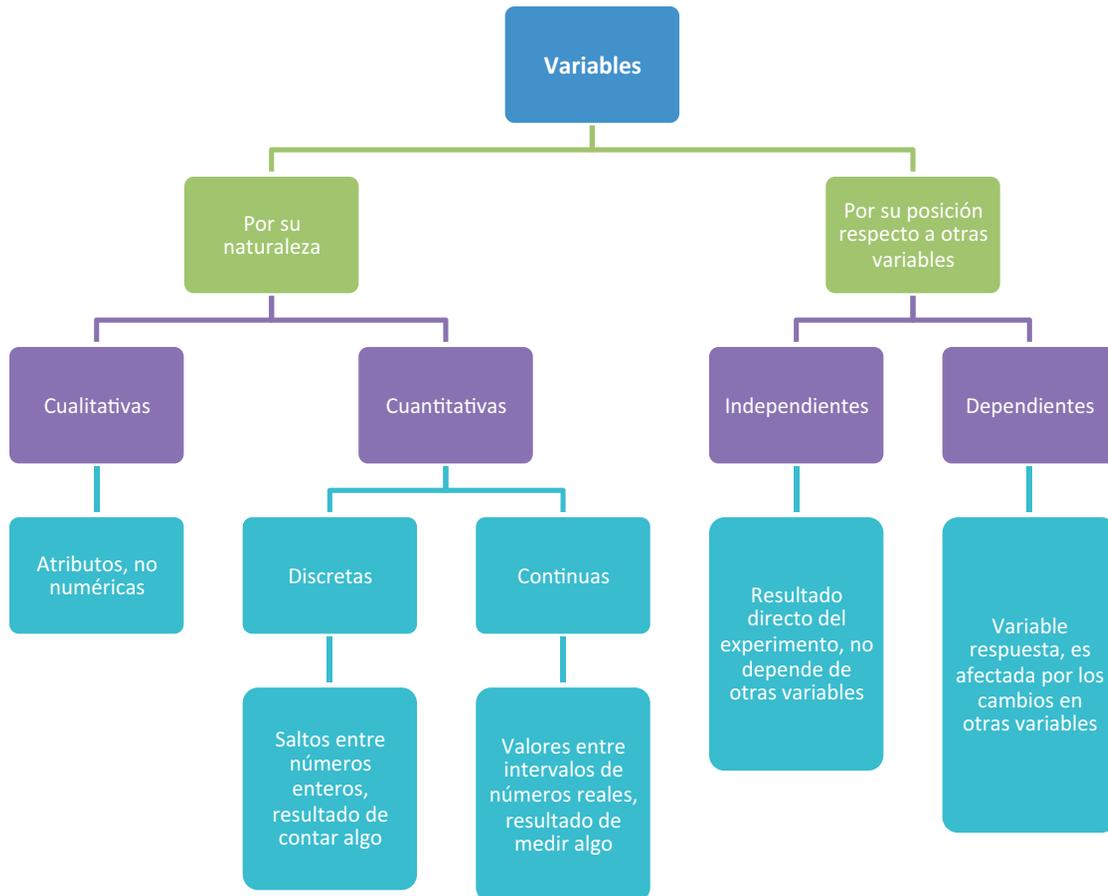
2. En el diagnóstico de tu comunidad, ¿estableciste variables dependientes e independientes? Menciona tres y argumenta tu respuesta.



Sabías que...

El concepto de “variabilidad” se refiere a los cambios que puede sufrir una variable.

Clasificación de las variables



Aprende más

Recolección de datos

Al iniciar este bloque dijimos que la Estadística se encarga de la recopilación, organización, análisis e interpretación de la información.

Como verás, la recolección de los datos es la primera de estas etapas. Ya hemos señalado que, para recolectar información puedes recurrir a fuentes directas o indirectas.

En el caso de las fuentes directas, hay diferentes herramientas para recolectar información. Recuerda que en la asignatura Taller de Desarrollo Comunitario que llevaste en tercer semestre conociste a fondo algunas técnicas de recolección de datos, como la observación, la entrevista y la encuesta.

Estos instrumentos te permiten medir las variables que deseas conocer. En el siguiente cuadro te presentamos un resumen de estas tres herramientas.



Reflexionemos sobre la actividad

¿De qué te das cuenta?

Antes de empezar el bloque 2 te sugerimos que analices la frase del filósofo francés Gastón Bachelard:

Lo importante no es medir para pensar, sino pensar para medir.

¿Qué quiere decir esto? Reflexiona y anota aquí tus conclusiones. Después discútelas con tus compañeros.

Cierre de bloque I

La Estadística te ayuda a dar significado a un conjunto de datos. En este bloque estudiamos la diferencia entre las variables discretas y continuas. Vimos que hay variables que se modifican por el comportamiento de otras y hay variables que no cambian por influencia de otras. A las primeras las llamamos variables dependientes y a las segundas variables independientes.

Además, aprendiste que la muestra es un subconjunto de la población y que su análisis (con ayuda de la Estadística descriptiva) permite, si la muestra está bien hecha, llegar a conclusiones sobre la población en general, a través de inferencia estadística.

En el siguiente bloque observarás que una buena recopilación de la información es indispensable para hacer buenos análisis y correctas interpretaciones estadísticas.

Autoevaluación

Lee detenidamente las preguntas y responde colocando una X en el nivel de avance que consideras que has logrado a lo largo del bloque I.

Interpretación del nivel de avance:

100-90% = Excelente, logré el aprendizaje de manera independiente.

89-70% = Bueno, requerí apoyo para construir mi aprendizaje.

69-50% = Regular, fue difícil el proceso de aprendizaje y lo logré parcialmente.

49% o menos = Insuficiente, no logré el aprendizaje.

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Conceptuales	Contenidos				
	Identificas para qué te sirve la estadística en tu vida diaria.				
	Comprendes la diferencia entre estadística descriptiva e inferencial.				
	Identificas los diferentes instrumentos de recolección de datos.				
	Reconoces en los textos diferentes tipos de variables.				

Bloque I

Comprendes y describes la variabilidad estadística y sus aplicaciones

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Procedimentales	Contenidos				
	Analizas críticamente la información que se te presenta.				
	Distingues los elementos de una población y una muestra.				
	Construyes ejemplos de población y muestra.				
	Explicas con tus propias palabras la diferencia entre variable discreta y continua, y variable dependiente e independiente.				

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Actitudinales	Contenidos				
	Valoras el trabajo en equipo como elemento que aporta y contrapone ideas en la resolución de problemas.				
	Cumples con las indicaciones dadas para el buen desarrollo de las actividades.				
	Buscas y sugieres soluciones a los problemas planteados.				
	Tienes una actitud positiva hacia el trabajo desarrollado en el bloque.				

BLOQUE II

Describe y representas datos,
tablas y gráficos



¿Qué aprenderás y cómo organizarás tu estudio?

Bloque II

12
HORAS

Objetos de aprendizaje que se abordan

- Organización de los datos.
 - ¿Para qué te sirve organizar la información?
 - Formas de organización de datos
 - Frecuencias: absoluta, relativa, acumulada y relativa acumulada
- Análisis de datos a través de tablas
 - Partes de una tabla
 - Tipos de tablas y su uso
- Análisis de datos a través de gráficas
 - Gráficas estadísticas
 - Partes de un gráfico
 - Histograma de frecuencia de clases
 - Polígono de frecuencias

Desempeños esperados al concluir el bloque

- Organizas y presentas los datos obtenidos en una distribución de frecuencias.
- Presentas una distribución de frecuencias en un histograma, un polígono de frecuencias y un polígono de frecuencias acumuladas.
- Construyes representaciones tabulares y gráficas después de reconocer el tipo de agrupación al que pertenecen los datos. Para obtener una mejor comprensión del comportamiento de la población del objeto de estudio.

Competencias disciplinares del campo de las matemáticas

- Organizas los datos en una distribución de frecuencias a partir de la variabilidad estadística observada y argumentarás su pertinencia.
- Expresas ideas y conceptos mediante representaciones lingüísticas, numéricas o gráficas relacionadas con la distribución de frecuencias.
- Construyes e interpretarás diferentes representaciones gráficas para la comprensión y análisis de situaciones.
- Formulas y resuelves problemas aplicando las reglas para agrupar datos en una distribución de frecuencias.
- Argumentas la solución obtenida a un problema, relacionándolo con la distribución de frecuencias.
- Operas las tecnologías de la información y la comunicación para obtener tablas de distribución de frecuencias y expresas las conclusiones de dicho proceso.

Productos de aprendizaje

- Elaboración del cuadro de conteo.
- Análisis de un conjunto de datos de tu comunidad.
- Elaboración de tabla de frecuencia absoluta y frecuencia absoluta acumulada.
- Construcción de la tabla de distribución de frecuencias de la comunidad de “El Capulín”.
- Elaboración de cuadro de distribución de frecuencias organizadas en clases de tu comunidad y la comunidad “El Capulín”.
- Análisis de datos a través de gráficas.
- Construyes el histograma y el polígono de frecuencias.
- Listado de las características de una muestra válida.

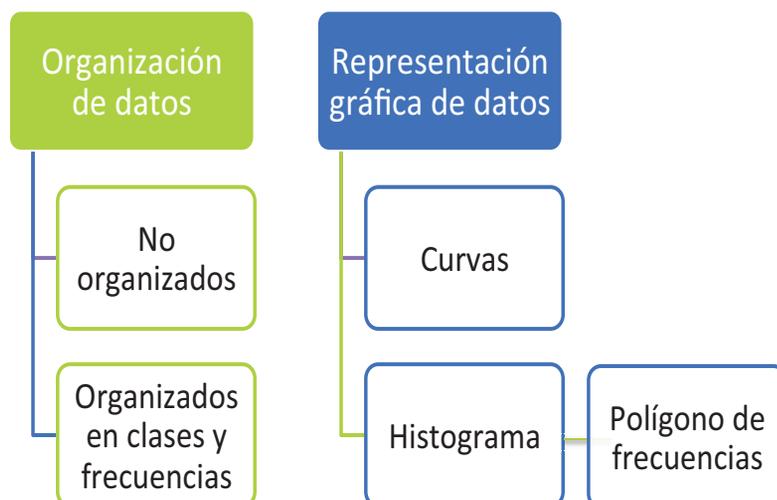
Introducción

En este bloque vas a aprender cómo recopilar información para identificar y analizar un problema. Vas a aprender también cómo organizar esta información para poder interpretarla. Asimismo, de qué manera puedes presentarla gráficamente una vez que la has organizado.

Por último, vas a saber qué es un histograma, qué es un polígono de frecuencias y, lo más importante, que más allá de ser representaciones visuales de una masa de datos, tienen propiedades estadísticas que te proporcionan mucha información valiosa.

Estos conocimientos te permitirán entrar de lleno en el mundo de la Estadística.

Mapa de objetos de aprendizaje





Aprende más

Organización de los datos

Existen muchas formas de organizar los datos. Para poder decidir cuál es la mejor debes tener muy claro para qué quieres la información. Recuerda a Bachelard:

Lo importante no es medir para pensar, sino pensar para medir.

¿Para qué te sirve tener organizada la información?

Imagínate que tu profesor te pide que hagas una presentación de tres minutos frente a tu grupo sobre tu comunidad.

¿Crees que puedes hablar de todo lo que hay en tu comunidad? Si no, ¿de qué hablarías? ¿Hablarías de su tamaño, de sus habitantes, de su vegetación y clima, de sus actividades productivas, de lo que haces como miembro de tu comunidad o hablarías de tu escuela, de lo que más te gusta, de lo que te gusta menos?

Toda esta información sobre tu comunidad es valiosa, pero ¿a qué características les darías prioridad al hablar? ¿Por qué? ¿Cómo la organizarías?



Supón que decides trabajar sobre las edades de los habitantes de tu comunidad y que su población es de 2,500 personas. Como no tienes el tiempo de ir a preguntarle su edad a cada uno de ellos, decides salir a la calle y entrevistar a las primeras 50 personas que te encuentras en el camino.

Los datos que encontraste son los siguientes:

Tabla 2.1

Edades de los 50 encuestados en la comunidad				
2	17	6	5	14
6	43	3	26	26
31	9	24	12	8
33	13	6	39	32
23	32	30	6	30
18	11	22	36	21
27	11	16	24	13
13	14	35	23	20
37	28	21	22	29
21	20	29	13	23



Actividad de aprendizaje 1

¿Qué observas en este conjunto de datos? Describe al menos cinco características que encuentres. Comparte tus reflexiones con tus compañeros.

1. _____
2. _____
3. _____
4. _____
5. _____

Formas de organización de datos

Tomemos los datos de edades de la sección anterior. ¿Cómo ordenarías la información?

No existe una forma única de ordenar los datos. Puede ser de menor a mayor, de mayor a menor, por decenas, entre otras. ¿Cuál es la mejor? La que te convenga, de acuerdo con el propósito de tu investigación.



Actividad de aprendizaje 2

Reúnete con tus compañeros en equipo y selecciona una de las formas de ordenar la información de edades en la siguiente tabla, la cual puede ser de: mayor a menor, menor a mayor, por grupos de edades repetidas (por columna). Una vez terminada la tabla de tu equipo, intercambien las tablas y discutan cuál de ellas permite conocer de manera más directa:

- ¿Cuántas personas están en el rango de edad de 20 a 30 años?
- ¿Cuál es el mayor grupo de personas con la misma edad?
- ¿Encuentras de forma más fácil las cinco características que describiste en la actividad anterior?

Comparte tus conclusiones con tus compañeros de otros equipos.

Tabla 2.2

Edades ordenadas de los 50 encuestados en la comunidad				

Tipos de frecuencia: absoluta, relativa, acumulada y relativa acumulada

Si te fijas únicamente en los resultados para las edades de 24, 26, 27 y 28 años observarás lo siguiente:

Tabla 2.3

Edad	Número de casos		Total
24			2
26			2
27			1
28			1

A este cuadro se le llama **cuadro de conteo**.

El cuadro de conteo te permite resumir la nube de datos. Ahora que sabes esto, elabora el cuadro de conteo de todos los datos de edades de las personas que entrevistaste.

Tabla 2.4

Cuadro de conteo de las edades de 50 habitantes de mi comunidad		
Edad	Número de casos	Total (f_a)

El cuadro de conteo te permite obtener información para organizar y facilitar el conteo de los datos. Al total de observaciones de cada renglón (edad en nuestro ejemplo) de tu cuadro de conteo se le llama frecuencia absoluta y se denota como *fa*. En tu cuadro puedes observar, por ejemplo, que la frecuencia absoluta para la edad de 26 es de dos, mientras que para la de 27 es sólo de uno. Como puedes observar, la columna de total del cuadro de conteo se llama **fat (frecuencia acumulada total)**.



Actividad de aprendizaje 3

A partir de tu cuadro, identifica los siguientes grupos de personas:

- a) La población que está cursando algún grado de primaria, considerando que, en condiciones óptimas, inicias la primaria a los 6 años y la concluyes a los 11.
- b) La población que está cursando algún grado de secundaria. En condiciones óptimas inicias la secundaria a los 12 años y terminas a los 14.
- c) La población en edad de votar.

1. ¿Cuántas personas hay en cada grupo?

Número de habitantes que:		
Cursan la primaria	Cursan la secundaria	Están en edad de votar



2. ¿Consideras que obtuviste del cuadro de conteo información más precisa de la que ya tenías sobre tu comunidad y sus habitantes? Argumenta tu respuesta.

3. Ahora supón que tu profesor te solicita que le presentes un breve informe al presidente municipal sobre el número de habitantes que pudieran recibir un apoyo económico del programa “Crece más”. Este programa del gobierno del estado atiende a las personas menores de 17 años. ¿Cómo obtendrías esta información?

4. ¿En qué fue diferente la forma en que obtuviste la información sobre el programa “Crece más” y sobre la población en edad de votar?



Aprende más

Análisis de datos a través de tablas

Seguramente te habrás dado cuenta de que para resolver el primer caso tuviste que sumar las frecuencias de la población de entre 2 y 16 años (los menores de 17 años). Para el segundo caso, las frecuencias de los que tienen 18 años o más. Sabiéndolo o no, en ambos casos utilizaste un nuevo concepto: **el de frecuencia absoluta acumulada (f_{aa})**.



Sabías que...

La frecuencia absoluta acumulada se obtiene de sumar la frecuencia absoluta de los datos anteriores con la faa del dato actual.



Actividad de aprendizaje 4

Como puedes observar en el cuadro siguiente, agregamos una nueva columna, la de frecuencia absoluta acumulada (f_{aa}).

1. Llena la columna de frecuencia absoluta acumulada, según como se indica en la tabla 2.6

Edades de 50 habitantes de mi comunidad		
Edad	Frecuencia absoluta (f_a)	Frecuencia absoluta acumulada (f_{aa})
2		
3		
5		
6		
8		
9		
11		
12		
13		

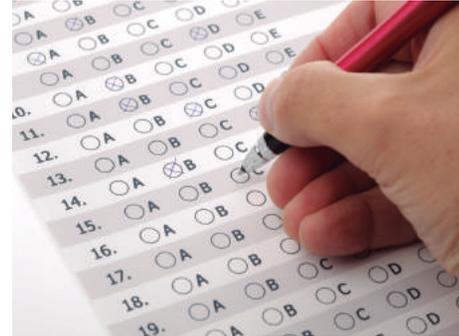
Tabla 2.7

Edades de 50 habitantes de mi comunidad			
Edad	Frecuencia absoluta (f_a)	Frecuencia absoluta acumulada (f_{aa})	f_{aa} en personas de 18 años o más
14			
16			
17			
18			
20			
21			
22			
23			
24			
26			
27			
28			
29			
30			
31			
32			
33			
35			
36			
37			
39			
43			
Total (fat)			

Bloque II

Describe y representas datos, tablas y gráficos

En el caso de la población en edad de votar, si hubieras tomado la frecuencia acumulada total y le hubieras restado la frecuencia acumulada hasta antes de la edad de 18 años tendrías el mismo resultado que obtuviste llenando la cuarta columna del cuadro.



En estos dos ejercicios resolvimos dos preguntas: en el caso del programa “Crece más” contestamos a la pregunta **menos de**, directamente con la tabla de frecuencias acumuladas. En el segundo caso, el de la población en edad de votar, resolvimos la pregunta **mayor de** con la fórmula:

$$\text{edad mayor de} = \text{edad total} - \text{observaciones acumuladas hasta antes del dato que buscamos.}$$

El informe que presentaste al presidente municipal le gustó. Por ello te ha pedido que realices el mismo ejercicio para los habitantes de la comunidad “El Capulín”. Te pide que compares los resultados que obtengas con los de tu comunidad.

Con la información que le proporciones piensa establecer el monto de recursos que van a utilizar en el programa “Crece más” en esa localidad. Como en el caso anterior te asignó un monto para que realices la investigación. Sin embargo, otra vez no te alcanzan ni el tiempo ni los recursos para entrevistar a todos los habitantes de “El Capulín”. Como la población de esta comunidad es cinco veces mayor que la tuya, decidiste encuestar cinco veces más personas que en el caso de tu comunidad, es decir 250 personas. Para hacer tu investigación tomaste el camión hasta la comunidad de “El Capulín”, te bajaste en la terminal y entrevistaste a las primeras 250 personas que te encontraste.

Los datos que obtuviste en la encuesta son los siguientes:

Tabla 2.8

Edades de 250 habitantes de "El Capulín"																								
51	7	89	74	53	69	22	83	7	11	75	83	17	24	60	61	87	84	63	8	7	52	31	48	84
74	70	3	37	35	12	52	29	30	49	3	75	19	5	77	63	80	37	37	69	90	48	50	45	10
1	83	87	20	55	5	63	59	38	27	84	8	79	31	36	61	833	38	88	26	79	17	51	86	9
72	82	30	88	69	19	43	52	44	76	48	73	12	76	85	7	28	5	48	6	68	13	79	81	89
8	8	17	78	13	43	76	86	4	28	84	55	81	40	11	78	83	45	32	67	50	18	78	2	88
85	65	11	63	14	81	86	76	3	25	44	21	19	68	70	41	28	63	65	84	14	45	80	44	76
68	37	65	16	33	20	10	81	32	38	76	76	49	71	53	6	32	17	19	86	8	43	77	27	7
73	61	62	67	4	44	28	11	88	88	9	64	90	56	10	21	34	53	37	89	37	5	54	73	84
48	71	15	33	35	50	56	71	21	21	9	90	36	60	72	43	47	16	50	17	42	58	57	81	44
80	51	27	66	50	89	13	62	44	44	64	19	50	5	34	84	88	23	4	60	26	20	77	53	86

2. ¿Qué semejanzas y diferencias encuentras al comparar ambas poblaciones: la de tu comunidad y la de “El Capulín”?

3. ¿Consideras que esta muestra de 250 personas tiene el mismo valor que el de las edades de las 50 personas que realizaste en tu comunidad? Argumenta tu respuesta.

4. ¿Cómo puedes comparar los resultados de dos comunidades de tamaños diferentes? ¿Tiene más valor hacer muestras más grandes?

La comparación de resultados en las dos comunidades de tamaño diferente la puedes hacer obteniendo para ambas los porcentajes de frecuencia de cada una de las edades registradas. Hacer esto se llama obtener las *frecuencias relativas*.



Actividad de aprendizaje 6

Completa la siguiente tabla de *frecuencias relativas* respecto al total de habitantes de la localidad “El Capulín”. Recuerda que fueron 250 personas.

1. A la tabla de tu comunidad agrégale dos columnas: una de frecuencia relativa y una de frecuencia relativa acumulada.
2. Haz lo mismo para la localidad “El Capulín”.
3. Compara tus resultados y comparte tus conclusiones con tus compañeros.

Recuerda: la frecuencia relativa es esencialmente las veces que aparece un dato repetido (f_a) dividido entre el número total de datos (N). Frecuencia relativa = f_a / N

4. Ahora responde las siguientes preguntas:

- ¿Cuál es el porcentaje de habitantes en edad de votar en ambas comunidades?
- ¿Cuál es el porcentaje de habitantes que son beneficiarios del programa “60 y + años” en ambas comunidades?

	Porcentaje de habitantes que:	
	Están en edad de votar	Son beneficiarios del programa “60 y + años”
Mi comunidad		
Comunidad “El Capulín”		

5. ¿Encuentras alguna similitud entre los datos obtenidos en ambas comunidades? Explica ampliamente:

6. ¿Qué ventajas tiene comparar en términos relativos las poblaciones de tu comunidad y de “El Capulín”?

- ---
- ---
- ---



Sabías que...

La frecuencia relativa se obtiene aplicando una regla de tres. Y la frecuencia relativa acumulada se obtiene sumando todas las frecuencias relativas.

Como ya dijimos, la frecuencia relativa es uno de los datos que te permite hacer comparables dos o más poblaciones de diferente tamaño que comparten una misma característica por ejemplo en nuestro ejercicio la edad de 84 años su frecuencia acumulada es de 6 ($250/100=6/x$).

Partes de una tabla

A la tabla que contiene la información de las frecuencias: absolutas, absolutas acumuladas, relativas y relativas acumuladas se conoce como **tabla de distribución de frecuencias**. Ahora bien, una misma información puede ser representada de diferentes maneras. Como hemos visto, una de ellas es a través de tablas.



Actividad de aprendizaje 7

Observa las siguientes tablas:

Tabla 2.11

OBESIDAD								
TABLA 1. PORCENTAJE DE LA POBLACIÓN MUNDIAL MAYOR DE 15 AÑOS CON UN ÍNDICE DE MASA CORPORAL SUPERIOR A 30.								
EEUU 31%	México 24%	Gran Bretaña 23%	Rusia 22%	Grecia 22%	Australia 22%	Nueva Zelanda 21%	Hungría 19%	República Checa 15%
Canadá 14%	España 13%	Irlanda 13%	Alemania 13%	Portugal 13%	Finlandia 13%	Turquía 12%	Bélgica 12%	Polonia 11%
Holanda 10%	Suecia 10%	Dinamarca 10%	Francia 9%	Austria 9%	Italia 9%	Noruega 9%	Japón 3%	Corea 3%

Fuente: OCDE 1996-2003

Disponible en http://www.censia.salud.gob.mx/descargas/infancia/2010/2.9._Estrategia_Salud_Alim.pdf
Consultado el 21 de abril de 2015.



Tabla de distribución de frecuencias: tabla donde se agrupan los datos en categorías mutuamente excluyentes que indican el número de observaciones en cada categoría.

Tabla 2.12

Sobrepeso y obesidad

Grupos	México		Distrito Federal	
	%	No. Personas	%	No. Personas
Mujeres mayores de 20 años	72	20.52 millones	75.4	2.3 millones
Hombres mayores de 20 años	66	16.96 millones	69.8	69.8 millones
Niños en edad escolar	26	5.54 millones	35	481,785 millones

Fuente: Olaiz-Fernández G, Rivera-Dommarco J, Shamah-Levy T, Rojas R, Villalpando-Hernández S, Hernández-Avila M, Sepúlveda-Amor J. ENSANUT 2006. Cuernavaca, México: INSP, 2006.
*Según datos de INEGI, 2005

Responde las siguientes preguntas:

1. ¿Qué información obtienes de cada tabla?

2. Identifica las variables en cada tabla y explica de qué tipo son.

3. ¿Qué elementos tienen en común ambas tablas?

4. Encierra en un círculo las frecuencias absolutas y las relativas.

5. ¿De dónde se obtuvo la información para hacer las dos tablas?

Como pudiste observar, toda tabla debe contener los siguientes elementos:



Título

Que identifica de qué se está hablando.



Valores numéricos

Información que permite al lector tener un panorama general sobre el conjunto de datos.



Renglones y columnas

Que dividen la información en categorías.



Fuente

Que refiere el origen de la información.



Actividad de aprendizaje 8

Una vez presentado tu informe, el presidente municipal te comenta que la información es demasiado detallada y que no le permite analizar lo que le interesa.

Te comenta que a él lo que le preocupa es la educación y que quiere saber: ¿cuántas niñas y niños aún no están en edad de estudiar la primaria; cuántos estudian la primaria, cuántos la secundaria; cuántos jóvenes tienen edad de estudiar el bachillerato y cuántos pueden cursar estudios superiores?

Ante esta situación, tú, recordando una vez más a Bachelard, reflexionas:

¿Cómo resolver esta petición? ¿Cómo dividir en grupos de edades tu población?
¿Cuántos grupos proponer? ¿Deben ser del mismo tamaño cada grupo?

Anota tus reflexiones y discútelas con tus compañeros.

Para empezar, elaboras un cuadro de distribución de frecuencias con las características que se muestran en el cuadro. Recuerda que la población mayor de 24 años estará representada en un solo grupo.

Tabla 2.13

Población en edad preescolar y en edad típica de estudiar en mi comunidad (Distribución de frecuencias organizadas en clases)							
Clases		Cuadro de conteo	Frecuencia absoluta f_a	Frecuencia absoluta acumulada f_{aa}	Frecuencia relativa f_r	Frecuencia relativa acumulada f_{ra}	Marca de clase m_i
Límite inferior	Límite superior						
0	menos de 6						
6	menos de 12						
12	menos de 15						
15	menos de 18						
18	menos de 24						
24	mayores de 24						
Suma / Total							

Con lo que has aprendido hasta este momento, ya puedes llenar todas las columnas, excepto la última, la de la **marca de clase**, que veremos un poco más adelante.

Cuando termines de llenar tu cuadro estarás en condiciones de responder al presidente municipal las siguientes dudas:



Clase: intervalos contruidos para agrupar la información.

1. ¿Qué proporción de la población todavía no está en edad de estudiar primaria? _____

2. ¿Qué proporción de la población está en edad de estudiar en educación básica (primaria+secundaria)? _____

3. ¿Qué proporción de la población está en edad de realizar estudios de bachillerato y estudios superiores? _____

4. ¿Qué proporción de la población no está en edad típica para estudiar? _____

Toda esta información, aunque es mucho más reducida que el total de datos que tenías inicialmente, le permitirá al presidente municipal tomar decisiones más precisas en materia de educación. ¡Felicidades!

Tu cuadro anterior contiene varios conceptos que son muy importantes:

- **Clase:** es cada uno de los grupos en los que concentras la información. (En nuestro ejemplo hay seis clases). Se denota como un intervalo.
- **Límite inferior de una clase:** es el valor de la variable con que inicia una clase.
- **Límite superior de una clase:** es el valor de la variable con que termina una clase.
- **Intervalo de clase:** es la distancia entre el valor del límite superior y el del límite inferior de una clase.
- **Marca de clase:** es el resultado de sumar los valores de los límites inferior y superior de cada clase y dividirlos entre dos. En pocas palabras, es el promedio de los valores de los límites inferior y superior de una clase.

Con esta información, ahora sí ya puedes llenar la columna de marca de clase.

Házlo y comparte tus conclusiones con tus compañeros.

Ahora que has terminado, te darás cuenta de que en tu cuadro puedes observar que:

- Tres de las clases tienen un mismo intervalo, de 0 a menos de 6 años, de 6 a menos de 12 años y de 18 a menos de 24 años.
- Dos de las clases tienen un intervalo de tres años: la que va de 12 a menos de 15 años y de 15 a menos de 18 años.
- Las cinco clases anteriores se llaman clases “cerradas”, pues tienen definido un límite inferior y uno superior.
- En cambio, la última clase que definimos es abierta, pues incluye las edades de 24 años o más, que en nuestro ejemplo llega a 43 años.

De acuerdo con Bachelard, definimos el número de clases, su tamaño (intervalo) y las características que interesan para nuestra investigación. Eso es lo correcto.

Sin embargo, en términos generales, algunos expertos proponen que es conveniente seguir, en la medida de lo posible, las siguientes reglas:

1. Hay que definir clases del mismo tamaño.
2. Puedes obtener el intervalo de clases restando la distancia entre el dato mayor y el dato menor de toda tu información y dividiendo ese resultado entre el número de clases que deseas formar.
3. La definición de los límites superior e inferior de una clase es distinta para variables discretas y variables continuas. En el primer caso, como la variable sólo toma valores “a saltos”, el cuadro podría quedar como sigue:

Tabla 2.14

Población que reporta número de años de estudio terminados		
Clases		
Límite inferior		Límite superior
0	-	6
7	-	12
13	-	15
16	-	21

En el caso de las variables continuas, el límite superior de una clase debe concluir en “menos de _____”, o el límite inferior en “más de _____”.



Sabías que...

Todo valor de tu conjunto de datos debe pertenecer exclusivamente a una clase. No puede haber un valor en más de una clase y no puede quedar fuera del cuadro ningún valor.



Actividad de aprendizaje 9

1. Completa la siguiente tabla de frecuencias para la comunidad “El Capulín”.

Tabla 2.15

Población en edad preescolar y en edad típica de estudiar para la comunidad “El Capulín” (Distribución de frecuencias organizadas en clases)								
Clases			Cuadro de conteo	Frecuencia absoluta f_a	Frecuencia absoluta acumulada f_{aa}	Frecuencia relativa f_r	Frecuencia relativa acumulada f_{ra}	Marca de clase m_i
Límite inferior	-	Límite superior						
0	-	menos de 6						
6	-	menos de 12						
12	-	menos de 15						
15	.	menos de 18						
18	-	menos de 24						
24	-	mayores de 24						
Suma / Total								

2. ¿Cómo puedes comparar la distribución de las edades de la población de tu comunidad y la de “El Capulín”?

Has escuchado la frase: “una imagen vale más que mil palabras” en este caso, una imagen vale más que mil números.

Como ya sabemos, no existe una forma única de representar los datos, pero el impacto visual de una gráfica permite sintetizar y explicar fácilmente lo que te interesa transmitir de la masa de datos.

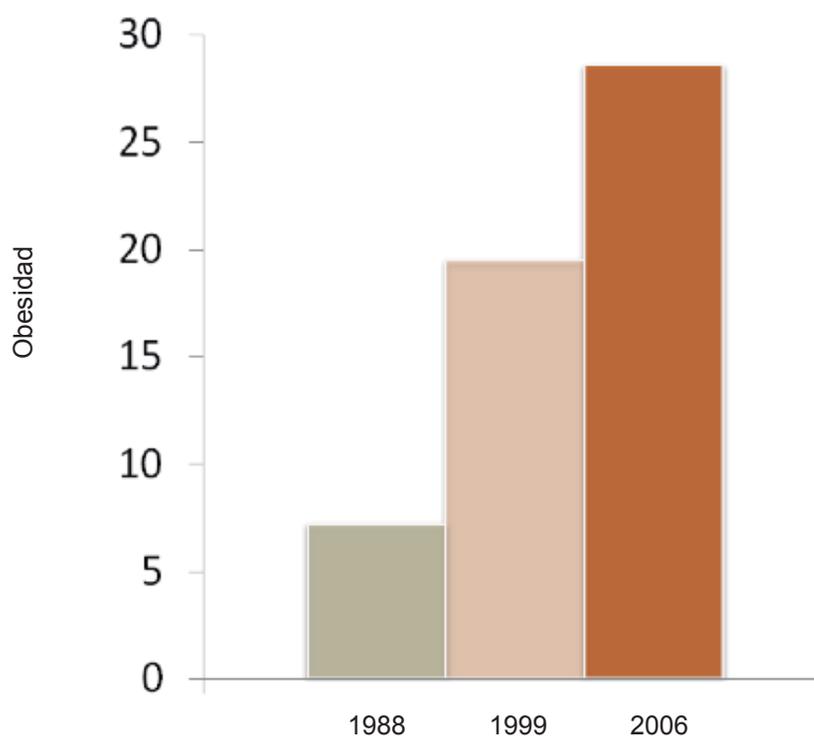
Tipos de tablas y su uso



Actividad de aprendizaje 10

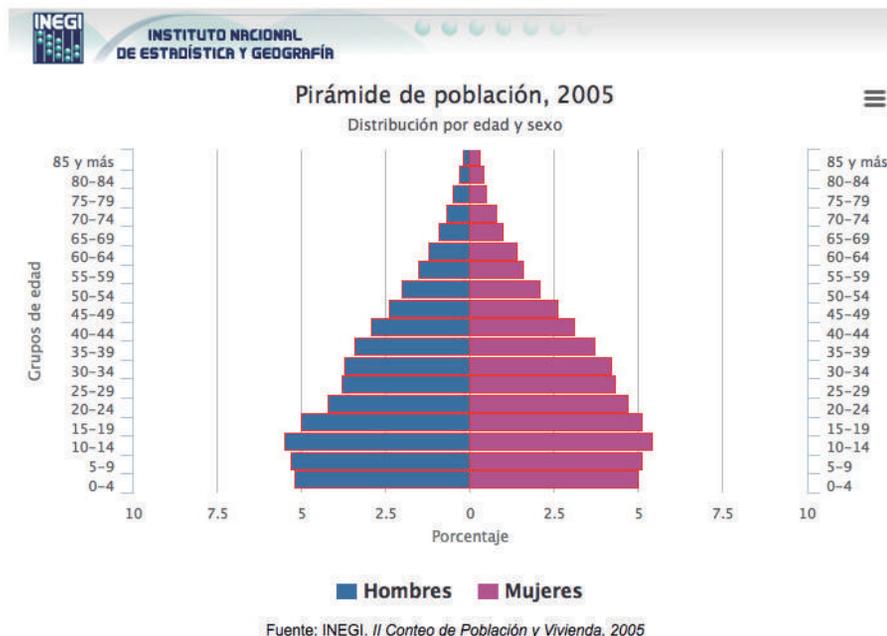
Observa las gráficas que te presentamos a continuación y responde:

Tabla 2.16 Aumento de la obesidad en México



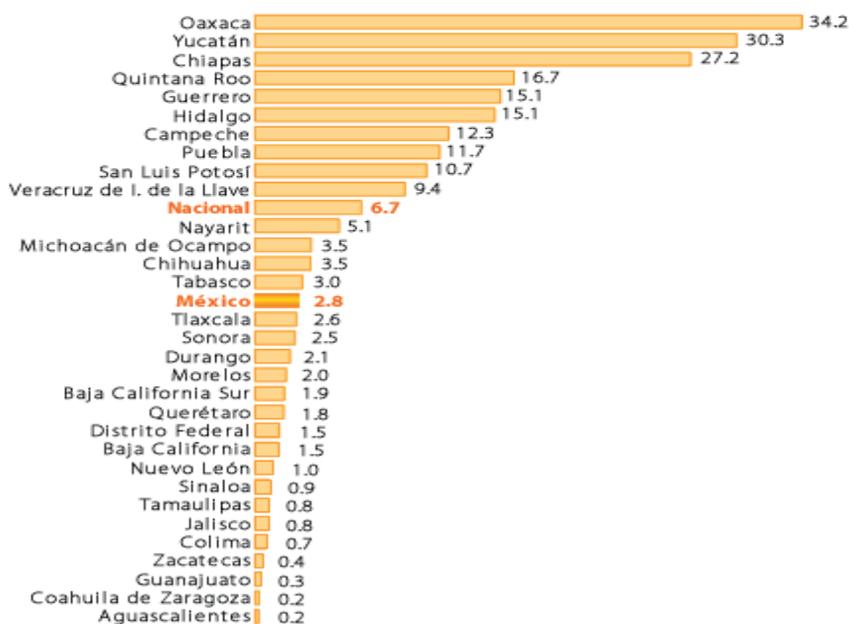
Disponible en: http://www.censia.salud.gob.mx/descargas/infancia/2010/2.9._Estrategia_Salud_Alimentaria.pdf Consultada el 21 de abril de 2015.

Tabla 2.17 Población total por miles por grupo de edad y sexo en 2008 y 2012 de un estado de la República Mexicana



Disponible en: <http://www3.inegi.org.mx/sistemas/temas/default.aspx?s=est&c=17484>
Consultado el 4 de junio de 2015.

Tabla 2.18 Porcentaje de la población que habla lengua indígena por entidad federativa (año 2010)



Disponible en: <http://cuentame.inegi.org.mx/monografias/informacion/mex/poblacion/diversidad.aspx?tema=me&e=15> Consultado el 4 de junio de 2015.

1. ¿Qué variables hay en cada una de las gráficas?

2. ¿Qué información te brinda cada una de éstas gráficas?

3. ¿Qué tipo de frecuencia utilizan en cada una de ellas?

4. ¿Qué similitudes y diferencias hay entre las gráficas?

Identifica los pasos que deberías seguir para obtener las mismas gráficas con la información de tu comunidad. No olvides que en tu Taller de Desarrollo Comunitario conociste diferentes tipos de instrumentos para la recolección de información.



Aprende más

Análisis de datos a través de gráficas

Dos gráficos muy utilizados en Estadística son:

**Histograma de
frecuencia de clases**

**Polígono
de frecuencias**

Histograma de frecuencia de clases

Como la primera de las gráficas anteriores, el **histograma** utiliza barras contiguas en donde la altura de cada barra corresponde a las frecuencias de clase. Sin embargo, existe una enorme diferencia entre ambas. Mientras que en la primera el ancho de cada barra no importa (sólo importa el concepto, que nos informa que se trata de los años 1988, 1999 y 2006); *en el caso de un histograma, sí es muy relevante.*

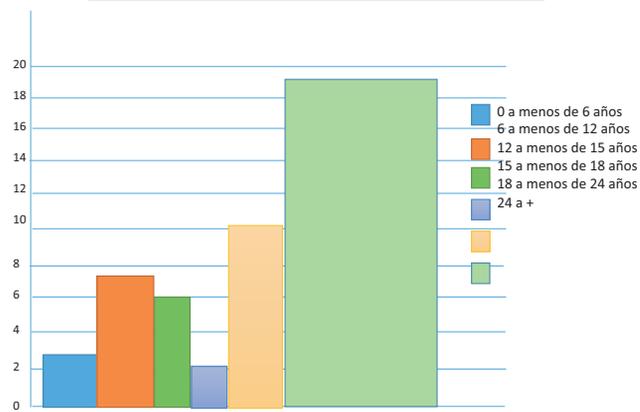
En el caso del histograma importan tanto el largo como el ancho de las barras o, en otras palabras, el área de cada una de las barras. Para poder entender por qué sí importa el ancho de cada barra de un histograma, te invitamos a que, a partir del cuadro de distribución de frecuencias de las edades de la población de tu comunidad construyas su **histograma**.



Histograma: representación gráfica de una distribución de frecuencias de una variable continua. Describe el comportamiento de un conjunto de datos en cuanto a su tendencia central, forma y dispersión. Está formado por un conjunto de rectángulos unidos, cuya base es igual a la amplitud del intervalo de clase y a la longitud proporcional a la frecuencia.

Tabla 2.19

Histograma de la distribución de edades de la población de tu comunidad por clases y frecuencias



Fuente: Datos recogidos por ti entre las primeras 50 personas que entrevistaste en tu comunidad

Polígono de frecuencias

El polígono de frecuencias está muy ligado al histograma, pues se construye trazando una línea poligonal formada por segmentos que unen los puntos medios de la base superior de las barras del histograma.

Como verás, en tu histograma las primeras dos columnas tienen un ancho de 6 años, las dos siguientes de 3 años; la quinta columna otra vez de 6 años y la última va de 24 a 43 años, es decir tiene un ancho de 19 años.

Ahora bien, si recuerdas de tus clases de matemáticas que la fórmula para calcular el área de un rectángulo es:

$$\text{Área de un rectángulo} = \text{base} \times \text{altura}$$

Entonces, el área de cada una de las barras del histograma será la misma, pues cada barra es un rectángulo.

Para saber la altura de la barra basta con ir al cuadro de conteo y observar la frecuencia de personas que caen en cada clase.

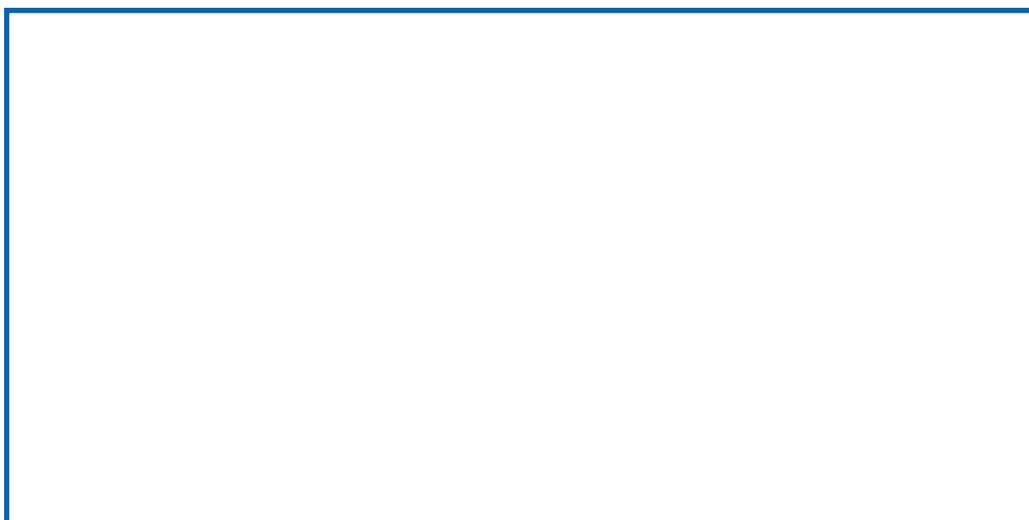
Sin embargo, saber el valor de lo ancho (la base) no es tan fácil.

- Si tomo el valor del límite inferior de la clase en cuestión estoy suponiendo que todas las frecuencias de esa clase caen en el límite inferior; en nuestro ejemplo, el de la clase que va de 0 a menos de 6 años, estaría suponiendo que todos caen en 0.
- Si, por el contrario, tomo el valor del límite superior, entonces, en nuestro ejemplo, estaría suponiendo que caen en 6.
- Para evitar estos sesgos extremos, se toma la marca de clase, que como ya sabemos, es el promedio aritmético de los valores de los límites superior e inferior.
- De esta manera, ya podemos calcular el área de cada clase, pues basta con multiplicar la frecuencia de una clase por la marca de clase que le corresponde.



Actividad de aprendizaje 11

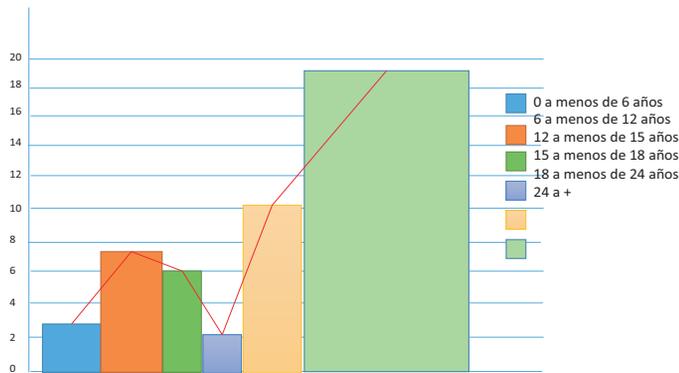
Calcula el área de cada una de las barras del histograma que construiste sobre las edades de la población de tu comunidad organizadas por clases y frecuencias:



El polígono de frecuencias se obtiene uniendo con segmentos de recta las marcas de clase de un histograma.

Tabla 2.20

Polígono de frecuencias de la distribución de edades de la población de tu comunidad por clases y frecuencias



Fuente: Datos recogidos por ti entre las primeras 50 personas que entrevistaste en tu comunidad

El área comprendida dentro del **polígono de frecuencias** equivale a la sumatoria de las frecuencias de todas las clases, multiplicadas por sus respectivas marcas de clase. Más adelante estudiaremos las propiedades que tienen, algunas curvas, específicamente la llamada curva normal. Estudiaremos también cómo estas curvas se utilizan para hacer inferencia estadística.

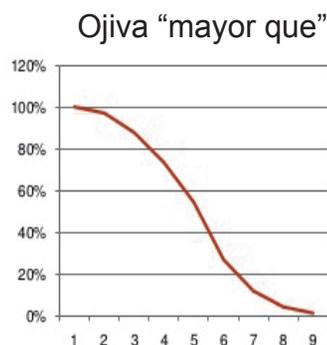


Polígono de frecuencias: gráfico utilizado para representar una distribución de frecuencias de una variable numérica, teniendo en cuenta su marca de clase.



Sabías que...

Otro tipo de curva se llama ojiva. Para trazar una ojiva utilizamos las frecuencias acumuladas. Con estas frecuencias se puede saber cuántos elementos de la distribución tienen un valor **“igual o menor que”** el límite superior de una clase elegida; o bien saber cuántos elementos del conjunto tienen un valor **“mayor que o igual”** al límite inferior de una clase elegida.





Actividad de aprendizaje 12

Antes de concluir este bloque te pedimos que hagas una reflexión final.

1. ¿La información que obtuviste hasta ahora de los habitantes de tu comunidad y la de “El Capulín”, es suficiente para generalizar sus resultados a cualquier otra comunidad de tu municipio? Explica ampliamente tu respuesta.

2. ¿Qué características debe tener una muestra para que sus resultados sean válidos para toda la población?



Sabías que...

Una población total se representa con una N , mientras que una muestra de esa población con una n .



Actividad de aprendizaje 13

Pon 5 ejemplos de ‘N’ y ‘n’. Comparte tus resultados con tus compañeros y con tu asesor.

N

n

Cierre de bloque II

La recopilación de la información te ayuda a analizar un problema, a conocer una población.

En este bloque estudiaste cómo, partiendo de una masa de datos, construimos un cuadro de conteo que nos ayuda a organizar la información para posteriormente resumirla en clases y construir la tabla de distribución de frecuencias.

Además aprendiste que hay cuatro tipos de frecuencias, las frecuencias absolutas, las frecuencias relativas y las frecuencias acumuladas, tanto absolutas como relativas.

Acuérdate de que las frecuencias relativas te permiten comparar dos poblaciones de tamaños diferentes y que las gráficas valen más que mil números, pues tienen un impacto visual que te permite sacar conclusiones rápidas de la masa de datos.

Por último, aprendiste a construir un histograma y un polígono de frecuencias y a diferenciarlos. Analizaste cómo el área de una distribución de frecuencias muestra características de una masa de datos que no se observan a simple vista.

Autoevaluación

Lee detenidamente las preguntas y responde colocando una X en el nivel de avance que consideras que has logrado a lo largo del bloque II.

Interpretación del nivel de avance:

100-90% = Excelente, logré el aprendizaje de manera independiente.

89-70% = Bueno, requerí apoyo para construir mi aprendizaje.

69-50% = Regular, fue difícil el proceso de aprendizaje y lo logré parcialmente.

49% o menos = Insuficiente, no logré el aprendizaje.

Describe y representas datos, tablas y gráficos

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Conceptuales	Contenidos				
	Organizas la información de acuerdo con el propósito de tu investigación.				
	Comprendes los conceptos de frecuencia absoluta, absoluta acumulada, relativa y relativa acumulada.				
	Identificas los elementos de un histograma y un polígono de frecuencias.				

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Procedimentales	Contenidos				
	Analizas críticamente la información que se te presenta.				
	Utilizas las tablas de distribución de frecuencias e identificas la función de cada una de sus partes.				
	Estructuras en clases un conjunto de datos de acuerdo con el propósito de tu investigación.				
	Representas gráficamente los datos de una tabla de distribución de frecuencias.				
Construyes y analizas las representaciones tabulares y gráficas de una población.					

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Actitudinales	Contenidos				
	Valoras el trabajo en equipo como elemento que aporta y contrapone ideas en la resolución de problemas.				
	Cumples con las indicaciones dadas para el buen desarrollo de las actividades.				
	Buscas y sugieres soluciones a los problemas planteados.				
Tienes una actitud positiva hacia el trabajo desarrollado en el bloque.					



BLOQUE III

Aplicas la estadística descriptiva



¿Qué aprenderás y cómo organizarás tu estudio?

Bloque III

16
HORAS

Objetos de aprendizaje que se abordan

- Estadígrafos de tendencia central.
 - Media aritmética
 - La mediana
 - Cuartiles, deciles y percentiles
 - La moda
- Estadígrafos de dispersión
 - Rangos
 - Desviaciones
- Algunas relaciones empíricas entre estadígrafos.
 - De tendencia central.
 - De dispersión.
- La curva normal.
- De la muestra a la población.

Desempeños esperados al concluir el bloque

- Calculas las medidas de tendencia central en diversas situaciones, a partir del conocimiento de los diferentes tipos de agrupación de datos para interpretarlos y analizarlos.
- Calculas las medidas de dispersión en diversas situaciones a partir del conocimiento de los diferentes tipos de agrupación de datos para interpretarlos y analizarlos.
- Interpretas el comportamiento de una población a partir de las medidas de tendencia central y dispersión de una muestra.

Competencias disciplinares del campo de las matemáticas

- Construyes, explicas e interpretas el valor de las distintas medidas de tendencia central y dispersión de una población.
- Eliges una medida de tendencia central o variabilidad, para la solución de un problema específico y argumentas su pertinencia.
- Manejas tecnología de información para obtener y expresar medidas de tendencia central o variabilidad en diversas situaciones.
- Formulas y resuelves problemas aplicando diferentes enfoques sobre las medidas de centralización o variabilidad para determinar algunas características de la población en estudio.
- Analizas las relaciones entre dos o más medidas de una población, para determinar su comportamiento.
- Estructuras argumentos acerca del comportamiento de una población, considerando sus medidas estadísticas.
- Interpretas el comportamiento de una población a partir de los resultados obtenidos utilizando tablas, gráficas, mapas, diagramas y textos con símbolos matemáticos y científicos.

Productos de aprendizaje

En este bloque realizarás los siguientes ejercicios para un conjunto de datos organizados por clases y frecuencias.

- Cálculo de la media aritmética.
- Cálculo de la mediana, deciles, cuartiles y percentiles.
- Cálculo de la moda.
- Cálculo de las desviaciones media, estándar y la varianza.
- La resolución de estas actividades formará parte de tu portafolio de evidencias.

Introducción

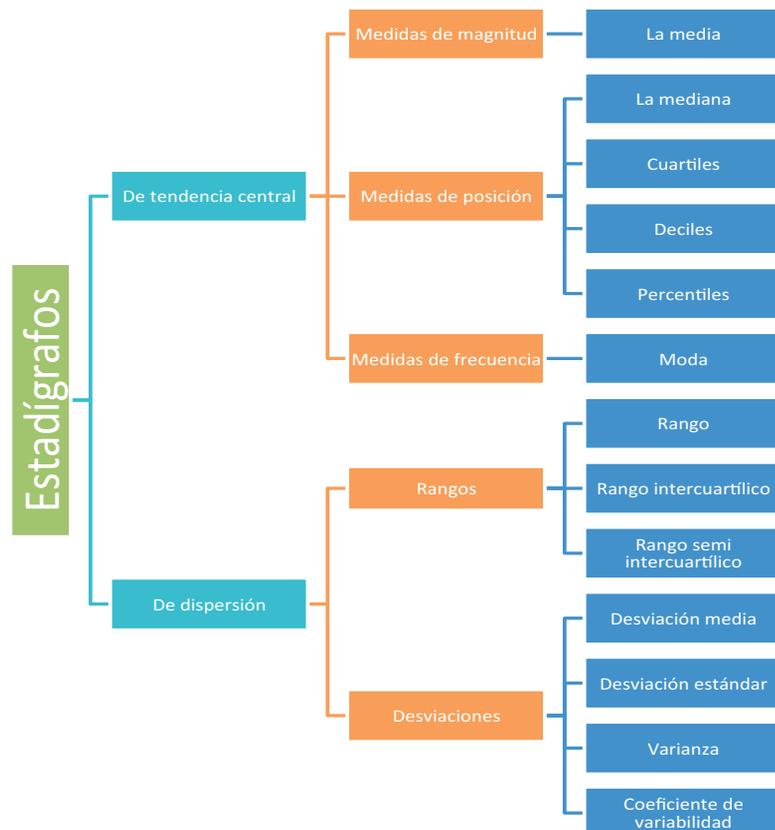
En el bloque anterior aprendiste a organizar información y a representarla gráficamente. Aprendiste también que, por ejemplo, en un histograma o polígono de frecuencias, puedes descubrir “patrones” de distribución que no se observan a primera vista. Para seguir profundizando en el análisis de un conjunto de datos, en este bloque vas a construir y manejar nuevos conceptos como: la media, los deciles y percentiles, la desviación estándar y la varianza. En términos generales, estos conceptos te van a permitir:

- a) Sintetizar toda la información recolectada en unos cuantos indicadores.
- b) Saber qué tan dispersa está una masa de datos.

A los indicadores que mencionamos en el inciso a) se les llama estadígrafos de tendencia central. En el inciso b) se incluyen los estadígrafos de dispersión. Con estos nuevos conceptos, y con lo que ya aprendiste en el bloque II, al terminar este bloque vas a saber:

- a) Organizar los datos en clases y frecuencias.
- b) Conocer el área que ocupa una masa de datos.
- c) Conocer algunas medidas representativas de esa masa de datos.
- d) Conocer qué tan representativas son las medidas mencionadas en el punto anterior.

Mapa de objetos de aprendizaje





Aprende más



La Estadística descriptiva te permite obtener conclusiones sobre un conjunto de datos y sólo sobre ese conjunto de datos (la muestra). De manera muy sencilla, la Estadística descriptiva incluye técnicas para recolectar, presentar, analizar e interpretar una masa de datos. En cambio, la Estadística inferencial incluye un conjunto de técnicas para obtener conclusiones que sobrepasan los límites de los conocimientos aportados por la muestra y busca explicar a partir de ellos cómo se comporta la población.

Al finalizar el bloque II escribiste cinco ejemplos sobre lo que es una población (N) y lo que es una muestra (n). A partir de ellos, vamos a iniciar este bloque señalando que la Estadística descriptiva tiene que ver con la muestra, mientras que la inferencia estadística se relaciona con la población.



Actividad de aprendizaje 1

Entender lo anterior es un paso importante en tu camino para ser un buen estadístico. Por ello, te invitamos a reflexionar:

1. ¿Qué es la Estadística descriptiva?
2. ¿Cuál es la diferencia entre la Estadística descriptiva y la inferencia estadística?
3. Argumenta tus respuestas y discútelas con tus compañeros y asesor.

A partir de tus respuestas te proponemos profundizar en tus conocimientos de Estadística descriptiva; es decir de la Estadística que estudia las características de una muestra.

Siguiendo con el ejercicio que has venido desarrollando desde el bloque anterior, supón que tu presidente municipal te pregunta cuál es la edad más representativa de las y los habitantes de tu comunidad. ¿Qué le contestarías?

- a) ¿Sumarías todas las edades y dividirías el resultado entre 50, que es el número de personas de tu comunidad que entrevistaste?
- b) ¿Ordenarías las edades de mayor a menor (o de menor a mayor) y escogerías la que quede exactamente en medio?
- c) ¿Le dirías que la edad más representativa es la que más veces se repite?; en el ejercicio que vienes realizando desde el bloque II, las edades de 6 y 13 años.
- d) Discútelo con tus compañeros y maestros.



Aprende más

Estadígrafos de tendencia central

Si seguiste el primer camino, le habrás sugerido al presidente municipal que la medida más representativa es la media aritmética. En cambio, si seguiste el segundo camino habrás hablado en favor de la mediana. Por último, si seguiste el tercero, le habrías recomendado que utilice la moda.

La media, la mediana y la moda son estadígrafos de tendencia central.

La media aritmética

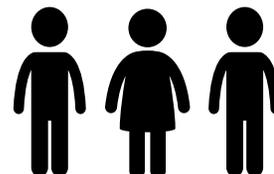
La media aritmética es el promedio que resulta de sumar el valor de todos los datos y dividirlo entre el número de datos.

En la media aritmética el valor de cada uno de los datos cuenta; esa es su principal ventaja. Por otro lado, tiene como principal desventaja que es muy sensible a los valores extremos. ¿Qué quiere decir eso?

Supón que te piden sacar el promedio de edad de Luis, Pepe y Laura. Cada uno de ellos tiene 15 años:

Luis	15 años
Pepe	15 años
Laura	15 años

Suman 45 años / 3 = 15 años



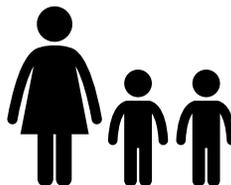
Bloque III

Aplicas la Estadística descriptiva

Supón ahora que te piden sacar el promedio de edad de:

Iván	5 años
Rafael	5 años
Lucía	35 años

Suman $45 \text{ años} / 3 = 15 \text{ años}$



Como puedes ver, en los dos casos la media aritmética es de 15 años, pero en el segundo la edad de Lucía eleva mucho el promedio. Por esto se dice que la media aritmética es sensible a valores extremos. Como se puede observar en el ejemplo, un solo dato con valor muy alejado del centro, aunque sea poco representativo por ser único, puede hacer variar significativamente el promedio.

La gran sensibilidad de la media aritmética a valores extremos tiene impactos importantes en las ciencias. Por ejemplo, tomó mucho a los economistas dar cuenta de que el ingreso por habitante (ingreso total del país entre el número de habitantes de ese país) no es un indicador suficiente de desarrollo económico, si existe una mala distribución del ingreso.

Otro ejemplo lo constituye la esperanza promedio de vida al nacer (edad promedio de vida de los habitantes de un país). Indicador que en México pasó de 39.8 años en 1940 a más de 75 actualmente. El drástico aumento de este indicador se explica por la importante disminución de la mortalidad infantil entre cero y cinco años.



Actividad de aprendizaje 2

Saca la media aritmética de las edades de los miembros de tu familia (considera únicamente las edades de tus papás y tus hermanos). Analiza si alguno o algunos de los valores sesgan la media aritmética. Argumenta tus respuestas.

La mediana

En la mediana, a diferencia de la media aritmética, no cuenta el valor de cada dato. Únicamente cuenta el valor de uno solo: el dato que divide la lista en dos mitades exactamente iguales.



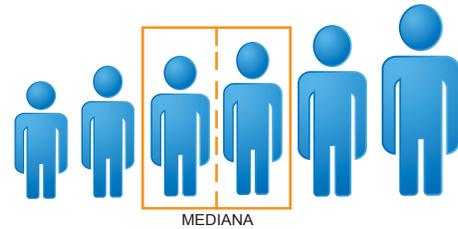
Sabías que...

El uso de la mediana en lugar de la media puede ser conveniente cuando existen valores extremos en un conjunto de datos.

Ejemplo 1:



Ejemplo 2:



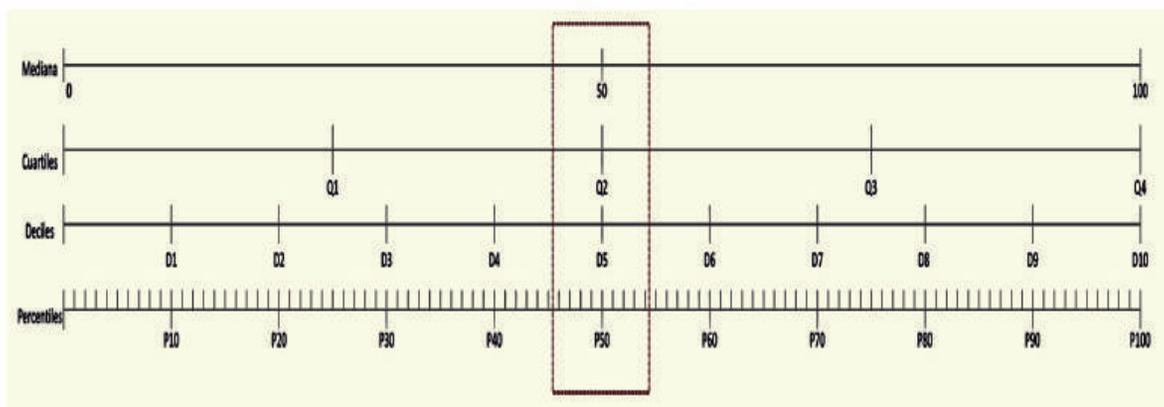
- En el ejemplo 1, como se trata de un número impar de cifras; la mediana será la persona que ocupa la posición 3.
- En el ejemplo 2, como se trata de un número par de cifras, la mediana será el promedio de edades de las personas que ocupan las dos posiciones contiguas.

Cuartiles, deciles y percentiles

Además de la mediana, existen otros estadígrafos que te permiten dividir tus datos ordenados en partes.

- Si divides tus datos en cuatro partes se llaman cuartiles. (Q)
- Si los divides en diez partes, deciles. (D)
- Si lo haces en cien partes, percentiles. (P)

Tabla 2.21 Estadígrafos de posición



Como puedes ver en el gráfico de arriba, la mediana, el segundo cuartil, el quinto decil y el 50 percentil, tienen el mismo valor.



Actividad de aprendizaje 3

Observa la gráfica anterior y responde:

1. ¿A qué equivalen en percentiles Q1, Q3, D3, D7 y P40?

2. En el gráfico siguiente encierra en un círculo la figura que corresponde a Q1 y Q3, D4 y D8.



Actividad de aprendizaje 4

Tabla 2.22

Universidad Regional del Noroeste
Exani-II
INFORME INDIVIDUAL PARA EL SUSTENTANTE

CONTROL: 0901001-1
Fecha: Mayo 2015

Nombre: _____

Resultados globales	%
Calificación (por ciento de aciertos obtenidos en la prueba)	82.6
Total de alumnos examinados	106

	0	25	50	75	100
Global	[Horizontal bar with markers at 0, 25, 50, 75, 100]				
Módulo I	[Horizontal bar with markers at 0, 25, 50, 75, 100]				
Módulo II	[Horizontal bar with markers at 0, 25, 50, 75, 100]				

COMPARACIÓN CON EL CONJUNTO DE ALUMNOS

= Inferior
 = Calificación
 = Superior

La imagen anterior muestra un reporte que te acaban de dar con el resumen de los resultados que obtuviste en el examen para ingresar a educación superior.

En ella se muestra una escala que divide en cuartiles los resultados de todos los estudiantes evaluados y posiciona los tuyos con respecto a ellos.

1. ¿Qué puedes decir de tu desempeño global y en los módulos I y II?

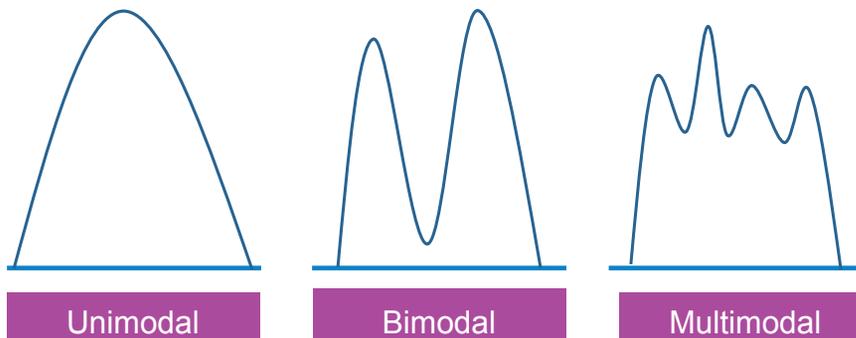
2. Marca con un círculo el Q1, Q2 y Q3.

3. Comenta con tus compañeros.

La moda

La moda es el valor que ocurre más frecuentemente en una distribución organizada de datos; es el valor que está de moda.

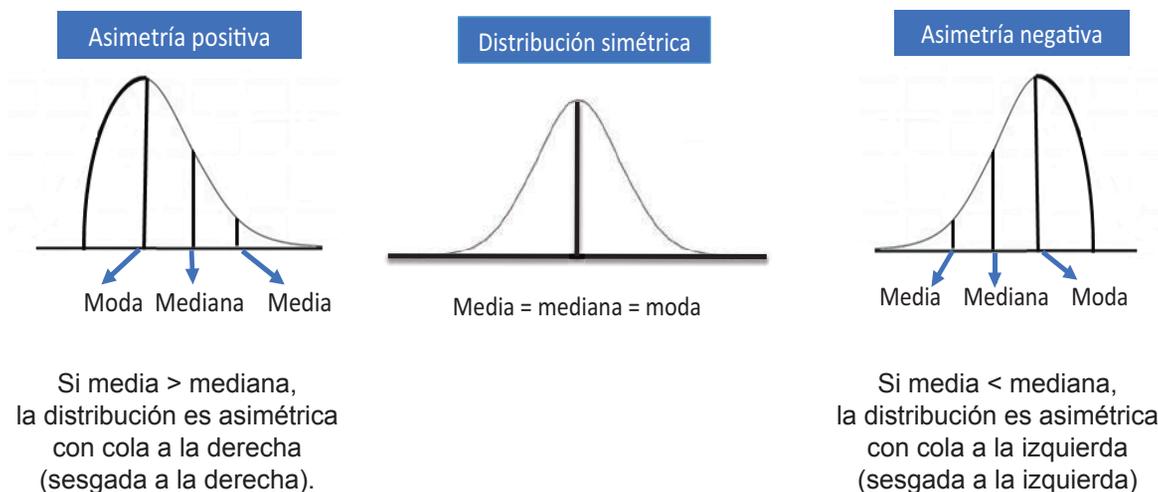
La moda puede no existir e incluso puede no ser única. Cuando hay dos modas en un conjunto de datos, se le llama conjunto bimodal; cuando hay más, multimodal.



A continuación te presentamos tres representaciones gráficas de la moda.

Relaciones empíricas entre la media, la mediana y la moda

En una distribución simétrica el valor de la media aritmética, la mediana y la moda coinciden.



Para curvas de frecuencia unimodales poco asimétricas existe la siguiente relación empírica: $Media - moda = 3 (media - mediana)$. Esto quiere decir que, en estos casos, la diferencia entre la media y la mediana es tres veces menor a la que existe entre la media y la moda.

Ya sabes qué es la media, la mediana y la moda. Ya conoces a algunos de los miembros de la familia de la mediana: los cuartiles, los deciles y los percentiles. Ya sabes cómo calcular todos estos estadígrafos si los datos no están agrupados.

¡Ahora viene lo bueno!

En las siguientes páginas vas a aprender a calcular todos estos estadígrafos cuando la información está organizada por clases y frecuencias.

Para ello, conviene recordar el cuadro de edades de tu comunidad, aunque con unos pequeños ajustes.

Tabla 2.23

Población en edad preescolar y los que idealmente no tienen edad para estudiar en mi comunidad (Distribución de frecuencias organizadas en clases)							
Clases	Número de observaciones	Marca de clase M_c	Intervalo	Frecuencia absoluta f_a	Frecuencia absoluta acumulada f_{aa}	Frecuencia relativa f_r	Frecuencia relativa acumulada f_{ra}
Los que no tienen edad de estudiar la primaria			0 a menos de 6				
Los que tienen edad para estudiar la primaria			6 a menos de 12				
Los que tienen edad de cursar la educación media (secundaria y bachillerato)			12 a menos de 18				
Los que están en edad de cursar la universidad			18 a menos de 24				
Los que no tienen edad para estudiar			24 a menos de 30				
			30 a menos de 36				
			36 o +				
Suma total							

¿Ya recordaste tu cuadro? ¡Qué bueno!

Si lo miras detenidamente observarás que hay dos cambios importantes:

En primer lugar, habrás observado que hay una nueva columna. La columna 1 que se llama “número de observaciones”.

En segundo lugar, que la marca de clase que en los cuadros del bloque II era la última columna, ahora es la número 2.

Estos pequeños cambios nos van a permitir:

- Ampliar el uso del concepto de marca de clase;
- Modificar la manera en que hemos entendido el concepto de frecuencia absoluta.

Respecto al primer punto, hasta ahora únicamente hemos utilizado la marca de clase para construir el polígono de frecuencias. Ahora lo vamos a utilizar para calcular de manera diferente la frecuencia absoluta, multiplicando la columna 1 por la 2 (marca de clase). Los resultados los vamos a anotar en la columna 3, “frecuencia absoluta”.

Seguramente te estarás preguntando, “pero ¿cómo? ¡Si antes llamábamos frecuencia absoluta a los datos que ahora aparecen en la columna 1, “número de observaciones”!

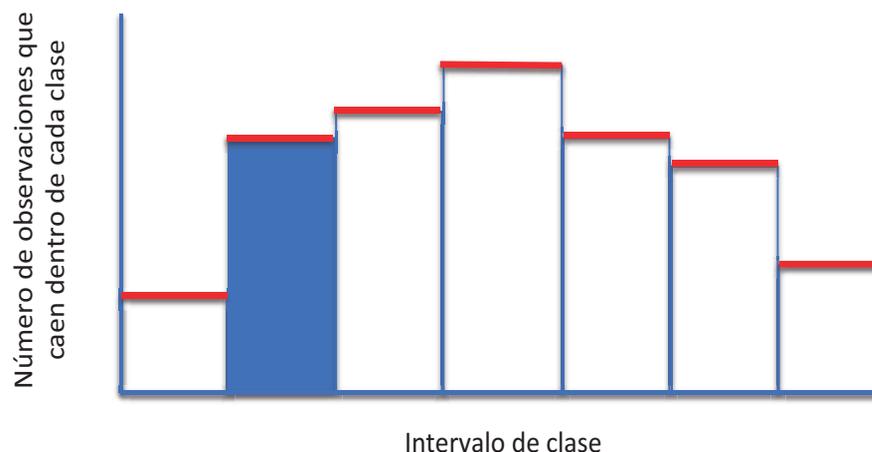
¡Es muy válida tu pregunta!

Lo que ocurre es que ahora la frecuencia absoluta es una frecuencia absoluta ponderada.

Para explicar este concepto, regresemos a nuestro histograma.

Tabla 2.24

Edades de la población de tu comunidad



En el histograma puedes observar que la primera definición de frecuencia absoluta únicamente considera la altura de cada una de las barras (marcada en rojo). En cambio, la segunda toma en cuenta toda el área de cada barra (marcada en azul).

Ahora sí, ¡pongamos manos a la obra!



Actividad de aprendizaje 5

A partir de lo que acabamos de comentar, calcula las frecuencias absoluta, absoluta acumulada, relativa y relativa acumulada de tu cuadro.

¡Sorpresa!

Cuando hayas concluido, tendrás a la vista los dos datos que necesitas para calcular la **media aritmética**:

$$\text{media aritmética} = \frac{\text{total de la columna 3}}{\text{total de la columna 1}} = \frac{(\text{total de frecuencias acumuladas})}{(\text{número total de datos})}$$



Actividad de aprendizaje 6

1. Calcula la media aritmética de las edades de las personas entrevistadas de tu comunidad, organizadas por clases y frecuencias.
2. Revisa si en el diagnóstico de tu comunidad que realizaste en el Taller de Desarrollo Comunitario de tercer y cuarto semestres calculaste algunas medias aritméticas.
3. Discútelas en grupo y explica la forma en que las utilizaste.



Sabías que...

La **media aritmética** de un conjunto de datos organizados en clases y frecuencias se obtiene:

1. Multiplicando las observaciones de cada clase por su marca de clase respectiva.
2. Sumando los resultados obtenidos para cada una de las clases de tu distribución de frecuencias.
3. Dividiendo el resultado entre el total de las observaciones de todas las clases.

Si la fórmula para calcular la **media aritmética de datos no organizados** era:

$$\bar{X} = \frac{\sum x_{ij}}{n}$$

Donde:

X_{ij} = valor de cada uno de los datos
 n = el número total de datos

La fórmula para calcular la **media aritmética de un conjunto de datos organizados** en clases y frecuencias es :

$$\bar{X} = \frac{\sum W_{ij} X_{ij}}{n}$$

Donde:

W_{ij} = total de observaciones de cada clase
 X_{ij} = la marca de cada clase
 n = es el número total de datos

Ahora veamos cómo se calcula la **mediana de un conjunto de datos organizados** en clases y frecuencias.

Lo primero que tienes que hacer es saber en qué clase cae la mediana. En tu ejercicio, como son 50 datos, (una serie par) la mediana estará entre los datos 25 y 26.

Si vas sumando las observaciones de las clases, verás que el dato 25 - 26 cae en la clase 4. (Ojo: tomas la columna de observaciones y no la frecuencia absoluta porque la mediana es un estadígrafo de tendencia central de posición, no importa su valor).

A partir de allí, lo demás es muy fácil. Usa la fórmula

$$Me = Li + \left[\frac{\left(\frac{n}{2}\right) - No. Obs\ acm}{No. Obs\ cm} \right] C$$

Donde:

Me = mediana

Li = valor del límite inferior de la clase en la que cae la mediana

n = número de observaciones

$No. Obs\ acm$ = número de observaciones acumuladas hasta antes de la clase mediana

$No. Obs\ cm$ = observaciones de la clase mediana

C = intervalo de clase

¿Muy fácil?



Clase mediana: clase a la que pertenece el valor de la mediana.

Ejemplo:

Tabla 2.25

Población en edad preescolar y en edad típica de estudiar en mi comunidad (Distribución de frecuencias organizadas en clases)							
Clases	Intervalo	Número de observaciones	Marca de clase M_c	Frecuencia absoluta f_a	Frecuencia absoluta acumulada f_{aa}	Frecuencia relativa f_r	Frecuencia relativa acumulada f_{ra}
Los que no tienen edad de estudiar la primaria	0 a menos de 6						
Los que tienen edad para estudiar la primaria	6 a menos de 12						
Los que tienen edad de cursar la educación media (secundaria y bachillerato)	12 a menos de 18						
Los que están en edad de cursar la universidad	18 a menos de 24						
Los que no tienen edad para estudiar	24 a menos de 30						
	30 a menos de 36						
	36 o más						
Suma total							

La Me cae aquí

No.Obs acm = número de observaciones acumuladas hasta antes de la clase mediana

No.Obs cm = observaciones de la clase mediana

Li = valor del límite inferior de la clase en la que cae la mediana.

n= número de observaciones

C= Intervalo de clase

Sólo ve la ecuación de abajo y sigue las flechas en el cuadro de arriba:

$$Me = Rojo + \left[\frac{\left(\frac{azul}{2} \right) - amarillo}{naranja} \right] verde$$



Actividad de aprendizaje 7

Resuelve el ejercicio.

¿Ves cómo sí era muy fácil?

¡Y lo mejor es que con la misma fórmula puedes resolver todos los casos de la familia de la mediana: los deciles, los cuartiles y los percentiles!

Desde luego que es necesario hacer cambios pequeños, pero muy pequeños.

¿Dónde está el dato del primer cuartil? Te pregunta una compañera que no trajo su libro de Estadística y Probabilidad I.

Como tú has seguido paso a paso este libro, le contestas: ¡Es muy fácil saberlo!

Como son 50 datos, y los cuartiles parten los datos en cuatro pedazos iguales; el primer cuartil será el dato 12.5; el segundo el dato 25 (idéntico a la mediana) y el tercero el 37.5.

Como buscas el primer cuartil, la fórmula será:
$$Q_1 = Li + \left[\frac{\left(\frac{n}{4}\right) - No. Obs ac Q_1}{No. Obs Q_1} \right] C$$

Donde:

Q_1 = primer cuartil

Li = límite inferior de la clase donde cae el primer cuartil

n = el número de datos

$No. Obs ac Q_1$ = número de observaciones acumuladas hasta antes de la clase del primer cuartil

$No. Obs. Q_1$ = número de observaciones de la clase en que cae el primer cuartil.

C = intervalo de clase

Como habrás visto, todo es idéntico al caso de la mediana, pero ahora referido al primer cuartil.

El único cambio que vale la pena comentar es que en la primera parte de la fórmula, en lugar de $n/2$ como fue para la mediana, ahora es $n/4$ (porque los cuartiles dividen la información en cuatro partes).

¿Y si hubiéramos estado buscando el segundo cuartil?

¡Excelente pregunta!

La fórmula sería:
$$Q_2 = Li + \left[\frac{\left(\frac{n}{4} \cdot 2\right) - No. de Obs ac Q_2}{No. Obs Q_2} \right] C$$

Como verás, tendrías que trabajar igual que para el primer cuartil, pero con los datos del segundo cuartil. El único cambio importante está encerrado en rojo en la fórmula de arriba:

Si recuerdas, para calcular el primer cuartil esta parte de la fórmula era $n/4$, que es igual a $n/4 \cdot 1$ (uno, porque se trata del primer cuartil).

Por lo tanto, para el segundo cuartil la fórmula sería $n/4 \cdot 2$. Para el tercer cuartil $n/4 \cdot 3$.

Con la misma lógica podrías encontrar el primer decil: si los deciles dividen la información en diez partes iguales y son 50 datos, el primer decil será el dato 5 y el cuarto decil el dato 20.



Actividad de aprendizaje 8

Calcula Q1, Q2 y Q3 con los datos de edades de los entrevistados de tu comunidad. Explica en palabras qué significan tus resultados.

Ya es obvio para ti que la fórmula para el decil 1 es:

$$D_1 = Li + \left[\frac{\left(\frac{n}{10} \cdot 1 \right) - No. Obs. ac D_1}{No. Obs_{1Decil}} \right] C$$

Y para el decil 6:

$$D_6 = Li + \left[\frac{\left(\frac{n}{10} \cdot 6 \right) - No. Obs. ac D_6}{No. de Obs_6Decil} \right] C$$

Y de la misma manera, para el percentil 23:

$$P_{23} = Li + \left[\frac{\left(\frac{n}{100} \cdot 23 \right) - No. Obs. ac P_{23}}{No. Obs P_{23}} \right] C$$

Y para el percentil 64

$$P_{64} = Li + \left[\frac{\left(\frac{n}{100} \cdot 64 \right) - No. Obs. ac P_{64}}{No. Obs P_{64}} \right] C$$



Actividad de aprendizaje 9

1. Calcula D_1 , D_6 , P_{23} y P_{64} .
2. En grupo calculen los cuartiles, deciles y percentiles que resulten pertinentes para fortalecer el diagnóstico comunitario que realizaste durante los dos semestres anteriores.



Sabías que...

- La media es un estadígrafo de tendencia central que mide magnitud.
- La mediana es un estadígrafo de tendencia central que mide posición.
- La moda es un estadígrafo de tendencia central que mide la frecuencia.

Vamos ahora a aprender a calcular la moda cuando los datos están organizados en clases y frecuencias. Lo primero que tienes que hacer es ubicar en qué clase está la moda. ¡Eso es muy fácil!

En tu cuadro, ve la columna de observaciones. La moda está en la clase que tiene la mayor cantidad.

Ahora sigue la fórmula que dice:

$$M^o = Li + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) C$$

Donde:

Li = es el límite inferior de la clase en la que cae la moda (mismo concepto que cuando calculaste la mediana).

Δ_1 = en cuánto es mayor el número de observaciones de la **clase modal**, respecto a la de la clase anterior.

Δ_2 = en cuánto es mayor el número de observaciones de la clase modal, respecto a la de la clase siguiente.

C = tamaño del intervalo de la clase modal (mismo concepto que para calcular la mediana).



Clase modal: es la clase en la que cae la moda.

Ejemplo:

Población en edad preescolar y en edad típica de estudiar en mi comunidad (Distribución de frecuencias organizadas en clases)							
Clases	Intervalo	Número de observaciones	Marca de clase M_c	Frecuencia absoluta f_a	Frecuencia absoluta acumulada f_{aa}	Frecuencia relativa f_r	Frecuencia relativa acumulada f_{ra}
Los que no tienen edad de estudiar la primaria	0 a menos de 6						
Los que tienen edad para estudiar la primaria	6 a menos de 12						
Los que tienen edad de cursar la educación media (secundaria y bachillerato)	12 a menos de 18						
Los que están en edad de cursar la universidad	18 a menos de 24						
Los que no tienen edad para estudiar	24 a menos de 30						
	30 a menos de 36						
	36 o más						
Suma total							

Li = límite inferior

C = Intervalo de clase



Actividad de aprendizaje 10

1. Calcula la moda de las edades de las 50 personas que entrevistaste en tu comunidad.
2. Explica tu respuesta y discútela con tus compañeros, compañeras y docentes.
3. Explica qué utilidad puede tener el cálculo de la moda en el diagnóstico de tu comunidad.



Aprende más

Estadígrafos de dispersión

Hasta aquí los estadígrafos de tendencia central. Ya conoces la media, la mediana y la moda, y sabes que son medidas que pueden representar a un conjunto de datos. Demos el siguiente paso:

Vamos a conocer ahora qué tan representativas son estas medidas de su conjunto de datos. Para ello se utilizan los estadígrafos de dispersión, que vamos ahora a trabajar.



Sabías que...

Los estadígrafos de dispersión pueden ser:

Rangos	Desviaciones
<input type="checkbox"/> Rango	<input type="checkbox"/> Desviación media
<input type="checkbox"/> Rango intercuartílico	<input type="checkbox"/> Desviación estándar
<input type="checkbox"/> Desviación intercuartílica	<input type="checkbox"/> Varianza

Como puedes ver en el cuadro, los primeros instrumentos que nos pueden ayudar a saber qué tan dispersa está la información con la que contamos son los rangos.

Rangos

El *rango* propiamente dicho mide la distancia que existe entre el mayor y el menor de los datos de un conjunto.

Dado que, en el cuadro de edades de tu comunidad, el mayor de los datos es el 43 y el menor el 2, el rango es igual a _____.

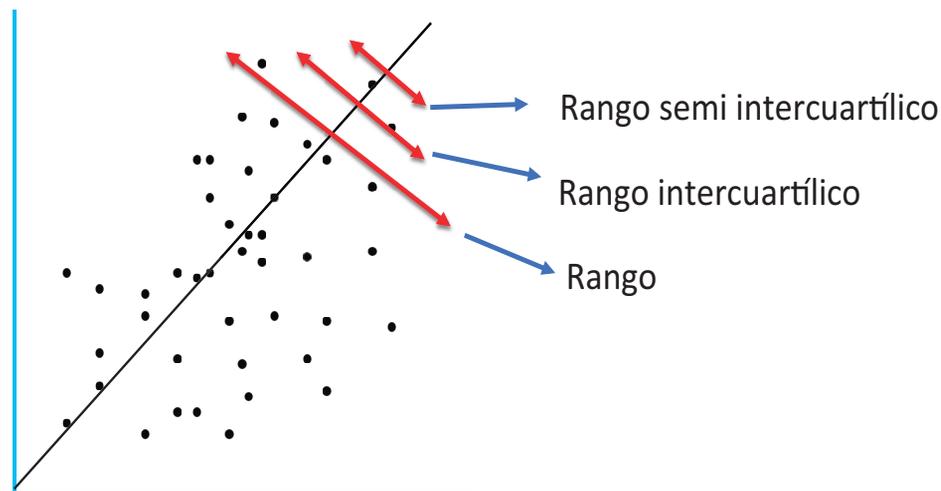
Por su parte, el *rango intercuartílico* mide la distancia entre el valor del cuartil 3 y del cuartil 1. Esto quiere decir que no considera la dispersión de todos los datos de un conjunto, sino únicamente la que existe entre los valores de Q3 y Q1.

Como tú ya calculaste los valores de Q3 y Q1 páginas arriba, te pedimos ahora que calcules rango intercuartílico = $Q_3 - Q_1$

$$Q_3 - Q_1 = \underline{\hspace{2cm}}$$

En el cuadro siguiente puedes apreciar la diferencia entre el rango y el rango intercuartílico:

Tabla 2.26 Conjunto de datos



Como puedes observar en el ejemplo de arriba, la ventaja del *rango* consiste en que incluye todas las observaciones. Sin embargo, su desventaja es que es muy sensible a la existencia de valores extremos.

En cambio, aunque el rango *intercuartílico* únicamente considera la mitad del trayecto entre el dato mayor y el dato menor, evitando el sesgo de los valores extremos.

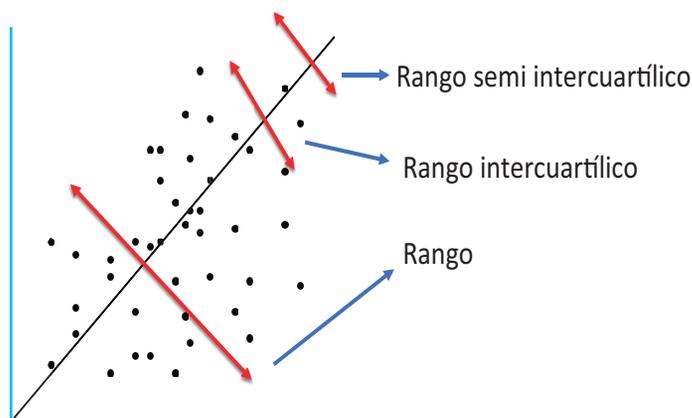
Existe un tercer rango que es aún más preciso: el rango semi intercuartílico. Si ves su fórmula lo vas a entender perfectamente.

$$\text{Rango semi intercuartílico} = \frac{Q_3 - Q_1}{2}$$

Como verás en el gráfico este rango considera la mitad de las observaciones que el rango intercuartílico.

Para interpretar correctamente los rangos intercuartílicos y semi intercuartílicos es necesario que recuerdes que los cuartiles son estadígrafos de posición. Esto quiere decir que pueden darse situaciones como la siguiente:

Tabla 2.27



En las que las distancias entre Q3 y Q1 no son simétricas alrededor de la media.

Desviaciones

Los rangos que hemos visto no utilizan todos los datos de un conjunto para medir la dispersión. Lejos de ello, utiliza sólo dos datos.

En el caso del rango propiamente dicho, únicamente los valores mayor y menor de un conjunto de datos; y en los casos de los rangos intercuartílico y semi intercuartílico los valores de Q3 y Q1.

En cambio, las tres desviaciones, como veremos a continuación: la desviación media, la desviación estándar y la varianza, se calculan comparando las diferencias de los valores *de todos y cada uno de los datos de un conjunto respecto la media aritmética* que tú conoces muy bien.

¡Sólo que tenemos un problema!

Como puedes ver en el siguiente gráfico, unos datos quedan por encima de la media y otros por debajo. Para cada dato por encima de la media, su diferencia respecto a ella será positiva. Y para el dato por debajo de la media; negativa.

El problema que tenemos es que si sumamos todas las diferencias positivas y les restamos las negativas, se compensan y dan cero para cualquier distribución de frecuencias.

Y entonces, obtener siempre cero como resultado no nos sirve de nada.

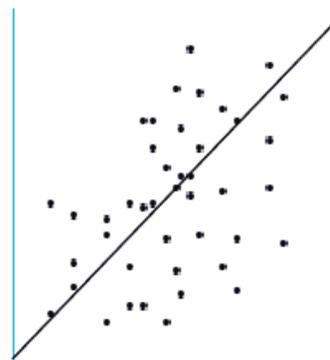


Tabla 2.28

De hecho, cada una de las tres desviaciones que mencionamos: la desviación media, la desviación estándar y la varianza, se distinguen por la manera en que resuelve este problema; ¡veámos cómo le hacen!

Ya vimos que la suma de las diferencias entre cada dato y la media siempre da cero.

En lenguaje matemático eso se escribe así:

$$\sum x_{ij} - \bar{x} = 0$$

Donde:

\sum = suma

x_{ij} = es cada dato de una distribución de frecuencias

\bar{x} = media aritmética

La primera manera de romper el cero es quitándole el signo a las diferencias entre cada dato y la media. Eso se llama manejar valores absolutos y se escribe así:

$$\sum |X_{ij} - \bar{x}| \neq 0$$

Donde:

| | = quiere decir valores absolutos y;

\neq = quiere decir diferente

Si el resultado obtenido lo dividimos entre el número de datos de un conjunto se llama **desviación media**:

$$Dm = \frac{|\sum ij - \bar{x}|}{n}$$

Donde:

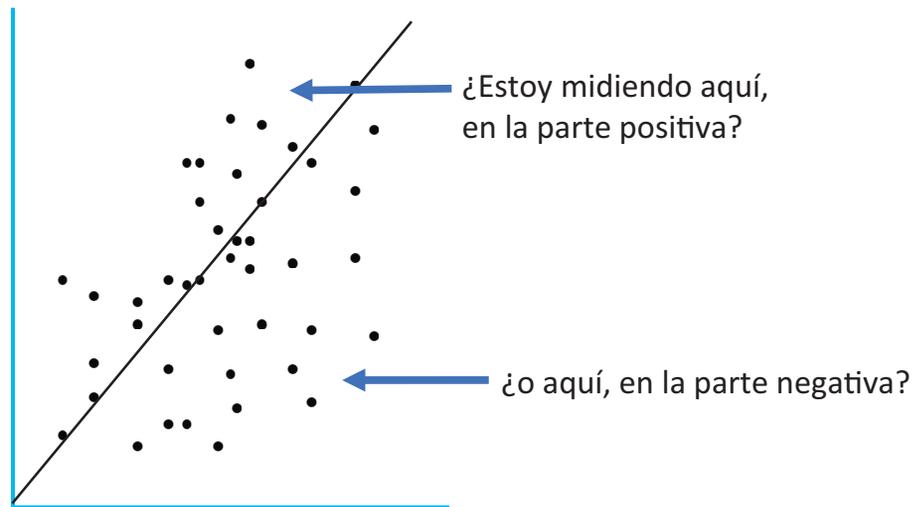
Dm = Desviación media

$\sum ij - \bar{x}$ = suma de las diferencias entre cada dato menos la media

n = el número de datos

Los enemigos de la desviación media la critican diciendo que al quitar los signos a la diferencia entre los datos y la media ya no se sabe si estamos midiendo la dispersión en la parte de arriba o de abajo de la distribución de frecuencias.

Tabla 2.29



Por lo tanto, hacen una segunda propuesta para romper el cero:

Como se eleva al cuadrado cualquier número negativo se vuelve positivo hay que elevar al cuadrado las diferencias de cada dato respecto a la media, y luego sacarle raíz cuadrada por eso estos críticos dicen que hay que sacarle raíz cuadrada para no alterar la ecuación.

$$\sqrt{\sum (X_{ij} - \bar{x})^2} \neq 0$$

Esta es la propuesta de la desviación estándar:

$$De = \sqrt{\frac{\sum (X_{ij} - \bar{x})^2}{n}}$$

Donde:

De = Desviación estándar
Lo demás = ya lo sabes

Es una buena solución, ¿no te parece? ¡A nosotros también!

Sin embargo, como siempre ocurre, hay quienes prefieren una tercera manera de romper el cero, porque no les gusta la raíz cuadrada de la fórmula anterior.

Para quitarla, proponen elevar al cuadrado los dos lados de la ecuación de la desviación estándar:

$$(De)^2 = \left(\sqrt{\Sigma(Xij - \bar{x})^2} \right)^2$$

que es igual a :

$$(De)^2 = \frac{\Sigma(Xij - \bar{x})^2}{n}$$

En donde:

$$(De)^2 = \text{varianza}$$

A la que llaman S^2 por sus siglas en inglés:

$$S^2 = \frac{\Sigma(Xij - \bar{x})^2}{n}$$



Actividad de aprendizaje 11

Calcula las desviaciones media, estándar y la varianza de los siguientes datos:
22, 9, 8, 13, 15, 24, 30, 14, 9, 3

Paso 1:

Calcula la media aritmética = _____

Paso 2:

Saca la diferencia de cada dato menos la media (tomando en cuenta el signo, la suma de todas las diferencias te debe dar cero).

Paso 3:

Para sacar la desviación media, suma, sin tomar en cuenta el signo, todas las diferencias y divide el total entre el número de datos.

Paso 4:

Para sacar la desviación estándar, eleva al cuadrado cada una de las diferencias de cada dato menos la media, suma todos estos resultados y el total divídelo entre el número de datos. A esa cantidad, sácale raíz cuadrada.

Paso 5:

Para sacar la varianza, haz todo lo del paso 4, excepto sacar la raíz cuadrada.

Bloque III

Aplicas la Estadística descriptiva

Ya sabes qué son: la desviación media, la desviación estándar y la varianza.

Ya sabes que, a diferencia de los rangos, las desviaciones utilizan todos los datos de un conjunto para medir la dispersión. Por lo tanto, son más confiables.

Ya sabes calcularlas para un conjunto de datos que no está organizado en clases y frecuencias.

Vamos ahora, por último, a aprender a calcular las tres desviaciones para un conjunto de datos organizados por clases y frecuencias.

Para ello, te invitamos a regresar a tu cuadro de edades de las 50 personas de tu comunidad que entrevistaste.

Tabla 2.30

		1	2	3	4	5	6	7	8	9	10
Clase	Intervalo	Obs	Mc	fa	faa	fr	fra	2-x	7*1	7 ²	9*1
1	0 a menos de 6										
2	6 a menos de 12										
3	12 a menos de 18										
4	18 a menos de 24										
5	24 a menos de 30										
6	30 a menos de 36										
7	36 a menos de 42										
Total											

Como recordarás, ya rellenaste toda la parte en azul.

Únicamente nos falta llenar las últimas cuatro columnas para calcular las desviaciones media, estándar y la varianza.

¡Pongamos manos a la obra!

Como ya hemos visto, lo primero que hay que hacer es calcular las diferencias de los datos respecto a la media. ¡Sólo que ahora los datos están organizados en clases y frecuencias!

Por eso, lo que antes hacíamos en un solo paso, restando a cada dato la media aritmética, ahora lo haremos en dos:

1. Restando a cada marca de clase la media aritmética.
2. Multiplicando cada uno de esos resultados por su respectivo número de observaciones.

Si miras tu cuadro, esas dos operaciones son las que te piden las columnas 7 y 8.

1. La columna 7 te pide restar la media aritmética a cada dato de la columna 2.
2. Por su parte, la columna 8 te solicita que multipliques cada resultado que obtuviste en la columna 7 por el dato que le corresponde de la columna 1.



Actividad de aprendizaje 12

1. Siguiendo las indicaciones anteriores, calcula los valores de las columnas 7 y 8 de tu cuadro.
2. Suma en el renglón de “total” los resultados de la columna 8.

¡Sorpresa!

Si hiciste tus cálculos, la suma total de la columna 8 te habrá dado cero.

¡Tú ya sabes por qué!

Si no tomas en cuenta el signo, la cantidad total que obtendrás, dividida entre el número de datos (los 50 entrevistados), te permite obtener la *desviación media* de un conjunto de datos organizados en clases y frecuencias.

Para sacar la desviación estándar, para datos no organizados, lo que tienes que hacer ahora es elevar al cuadrado las diferencias entre cada dato y la media. Sólo que ahora, este cálculo lo tienes que hacer también en dos pasos.

Como seguramente ya sospechas, estos dos pasos los resuelves con las columnas 9 y 10 y te van a permitir calcular la *desviación estándar* y la *varianza*.

Primer paso: en la columna 9 elevas al cuadrado la diferencia de cada marca de clase respecto a la media.

Segundo paso: en la columna 10, multiplicas cada uno de los resultados del primer paso por el número de observaciones que le corresponde (columna 1).

Cuando hayas dado estos dos pasos, únicamente deberás sumar todos los resultados de la columna 10, la raíz cuadrada de ese total dividido entre 50 (el número de personas de tu comunidad que entrevistaste) es igual a la *desviación estándar*.

Y, como ya sabes, la *varianza* la obtienes de manera aún más sencilla. Basta con hacer la suma de todos los resultados de la columna 10 y dividir ese total entre 50. Esa es la varianza.



Actividad de aprendizaje 13

Como ya tienes todas las herramientas, calcula las desviaciones media, estándar y la varianza de tu cuadro de edades de la población de tu comunidad.

Recordando una vez más a Bachelard, analiza tus resultados y discútelos con tus compañeros y maestros.





Sabías que...

1. La fórmula para calcular la desviación media:

a) Para datos no organizados es:

$$Dm = \frac{\sum |X_{ij} - \bar{x}|}{n}$$

Donde:

Dm = Desviación media

\sum = Suma

$| |$ = en valores absolutos

X_{ij} = cada uno de los datos de un conjunto

n = el número de datos

b) Para datos organizados en clases y frecuencias:

$$Dm = \frac{\sum f_{ij} |X_{ij} - \bar{x}|}{n}$$

2. La fórmula de la desviación estándar:

a) Para datos no organizados:

$$De = \frac{(\sqrt{\sum (X_{ij} - \bar{x})^2})^2}{n}$$

Donde ya conoces todos los símbolos.

b) Para datos organizados en clases y frecuencias:

$$De = \sqrt{\frac{\sum f_{ij} (X_{ij} - \bar{x})^2}{n}}$$

Donde ya conoces todos los símbolos.

3. La fórmula de la varianza:

a) Para datos no organizados es:

$$De^2 = \frac{\sum(X_{ij} - \bar{x})^2}{n}$$

b) Para datos organizados en clases y frecuencias:

$$De^2 = \frac{\sum f_{ij}(X_{ij} - \bar{x})^2}{n}$$

Donde ya conoces todos los símbolos.

La desviación estándar y la varianza están muy relacionadas. La primera es la raíz cuadrada de la segunda.

En cambio, el rango intercuartílico, la desviación media y la desviación estándar difieren de tamaño. Para desviaciones normales (platicaremos algo sobre estas distribuciones antes de cerrar este bloque).

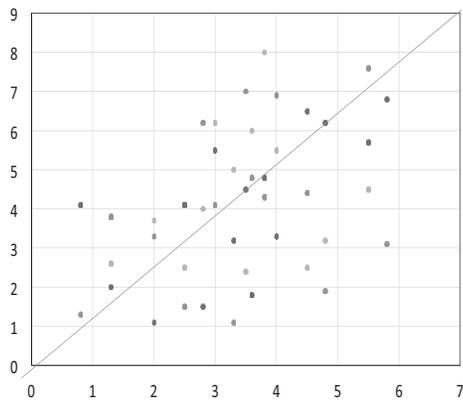
$$Dm = \frac{4}{5} De = 0.7979$$

$$Rango IQ = \frac{2}{3} De = 0.6745$$

Sólo en el caso de que todos los datos cayeran sobre la media, todos los rangos y las desviaciones serían iguales, y su valor sería cero.

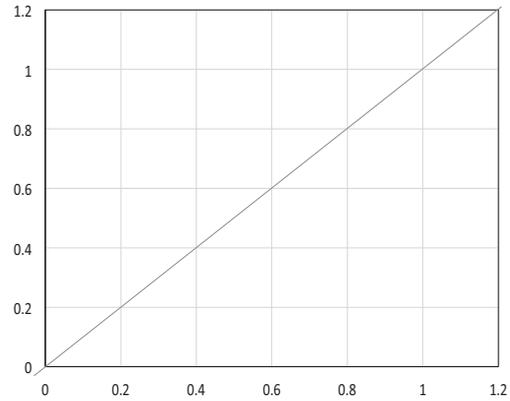
Ejemplo:

Tabla 2.31



Caso en el que **NO** ocurre que todos los datos caen en la media

Tabla 2.32



Caso en el que **SÍ** ocurre que todos los datos caen en la media



Sabías que...

Hoy existen varios programas y paquetes de computación que te permiten realizar estos cálculos estadísticos. Entre los más comunes están: las hojas de cálculo de Excel o el SPSS. El coeficiente de variabilidad es el resultado de dividir la desviación estándar entre la media aritmética.

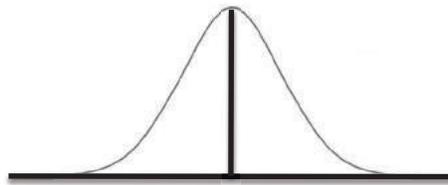
Mide qué tanto peso tienen la desviación de los datos de una distribución respecto a la media aritmética. A mayor coeficiente de variabilidad, mayor es el peso.

Cierre de bloque III

Reflexiona sobre lo aprendido

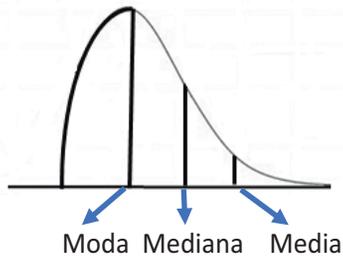
Al finalizar el bloque II supiste que el histograma y el polígono de frecuencias permiten conocer el área (la forma) de un conjunto de datos.

Distribución simétrica



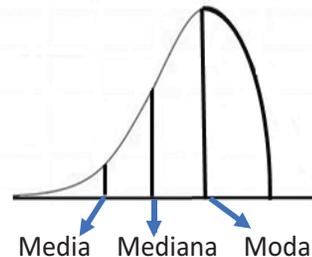
Media = mediana = moda

Asimetría positiva



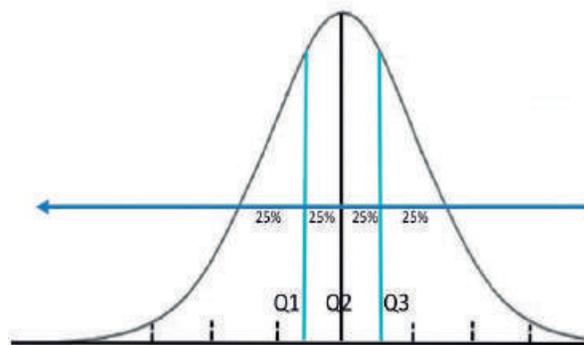
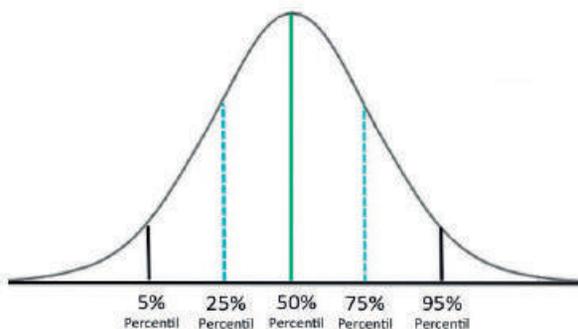
Moda Mediana Media

Asimetría negativa

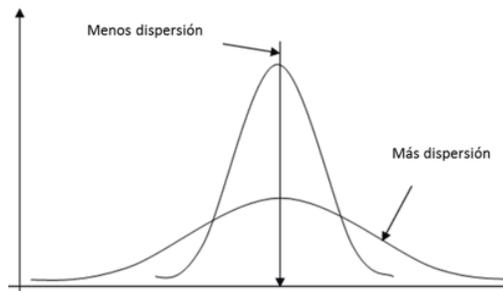


Media Mediana Moda

En este bloque III aprendiste que hay algunas medidas (estadígrafos de tendencia central y algunos parientes de la mediana, como los deciles y los cuartiles) que pueden representar a un conjunto de datos.

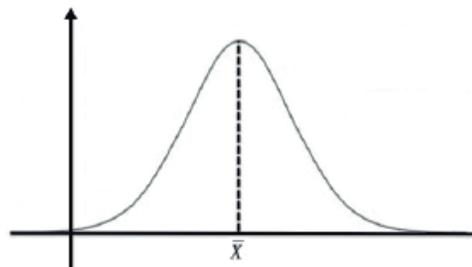


Aprendiste también que hay dos formas de saber qué tan representativas son las medidas incluidas en el punto anterior: los rangos y las desviaciones. Ahora ya sabes que mientras los rangos únicamente utilizan un par de valores para medir la dispersión, las desviaciones utilizan todos los datos, por lo que son más precisas.



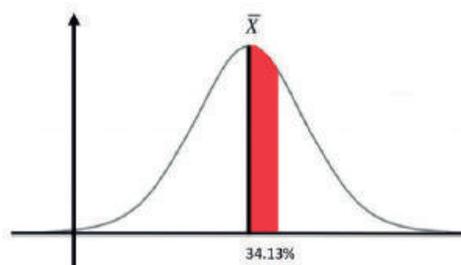
Para concluir este bloque te queremos presentar una curva que ha sido muy estudiada: la curva normal.

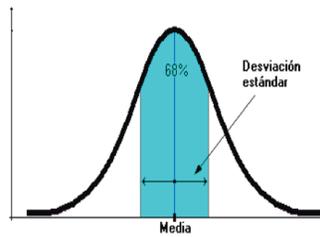
Esta curva tiene como característica principal el hecho de que muchos fenómenos sociales, físicos, químicos y económicos tienen su forma, es decir, se distribuyen como una curva normal, por ejemplo la estatura de la población, el efecto de una medicina en un grupo de enfermos, el tiempo que duran prendidos los focos que han sido producidos con materiales y en condiciones similares.



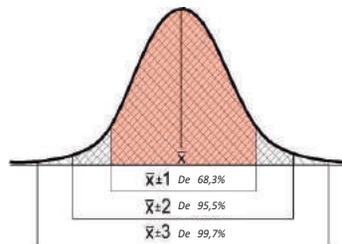
Como puedes observar, la curva normal es simétrica, tiene un solo pico central y media aritmética, mediana y moda son iguales.

Otra característica de esta curva es que se sabe que 34.13% de los casos están entre la media y una desviación estándar. Esto quiere decir que 68.26% de las observaciones caen a una desviación estándar de cada lado de la media.





Quienes han estudiado esta curva han demostrado que 95% de todos los casos caen entre la media y dos desviaciones estándar a cada lado de la distribución. Y que 99.3% de las observaciones quedan entre la media y tres desviaciones estándar.



¿Y para qué sirve todo esto?, te preguntas.

Pues, aunque a simple vista no lo parezca, esto es de la mayor importancia; tanta que ha permitido muy importantes avances científicos y en la vida de la población.

Supón que la distribución de las edades de los 50 entrevistados de tu comunidad te dio una curva normal, con media de 20.7 y desviación estándar de 1.44.

Supón que la muestra con la que has venido trabajando es una muestra representativa (aleatoria)

Si estos dos supuestos se cumplieran, podrías ir a ver al presidente municipal de tu comunidad y presumir diciéndole:

“Señor presidente municipal, le puedo asegurar, que con un nivel de confianza de 95%, que si usted escoge una persona de nuestra comunidad al azar, su edad estará entre los 19.3 y los 23.5 años de edad”.

Nota: Desafortunadamente el segundo supuesto no se cumple.

¿Recuerdas que cuando empezaste tu ejercicio dijimos que para entrevistar a la gente de tu comunidad saliste de tu casa y entrevistaste a las primeras 50 personas que te encontraste?

¿Recuerdas también que cuando fuiste a hacer la encuesta a la comunidad de “El Capulín” te bajaste en la estación del camión y entrevistaste a las primeras 250 personas que se te atravesaron?

Pues ¡malas noticias!

Ninguna de esas dos formas de entrevistar te permitieron hacer una muestra aleatoria, por lo que tus conclusiones sólo valen para las 50 personas que entrevistaste en tu comunidad y las 250 de la comunidad “El Capulín”, no para toda la población de tu comunidad, ni para toda la población de “El Capulín”.

En otras palabras, ya puedes analizar una muestra, pero todavía no puedes sacar inferencias de esa muestra para toda la población.

Pero esto no debe desanimarte. Lo que pasa es que ya sabes Estadística descriptiva... pero todavía no conoces la inferencia estadística.

Ahora supón que ya pasaron los años desde que estudiaste el telebachillerato comunitario, que eres médico y sabes hacer una muestra aleatoria y acabas de probar una vacuna en una muestra que tú definiste.

Lo más probable entonces es que el resultado sea una curva normal y que puedas afirmar con orgullo:

“Señor secretario de salud, con un nivel de confianza de 95%, le puedo asegurar que esta vacuna es la cura para **todos** los casos de la enfermedad que nos venía afectando”.

Para avanzar a que esto sea posible, pasemos ahora al bloque IV para adentrarnos en el apasionante mundo de la Probabilidad.

Autoevaluación

Lee detenidamente las preguntas y responde colocando una X en el nivel de avance que consideras que has logrado a lo largo del bloque III.

Interpretación del nivel de avance:

100-90% = Excelente, logré el aprendizaje de manera independiente.

89-70% = Bueno, requerí apoyo para construir mi aprendizaje.

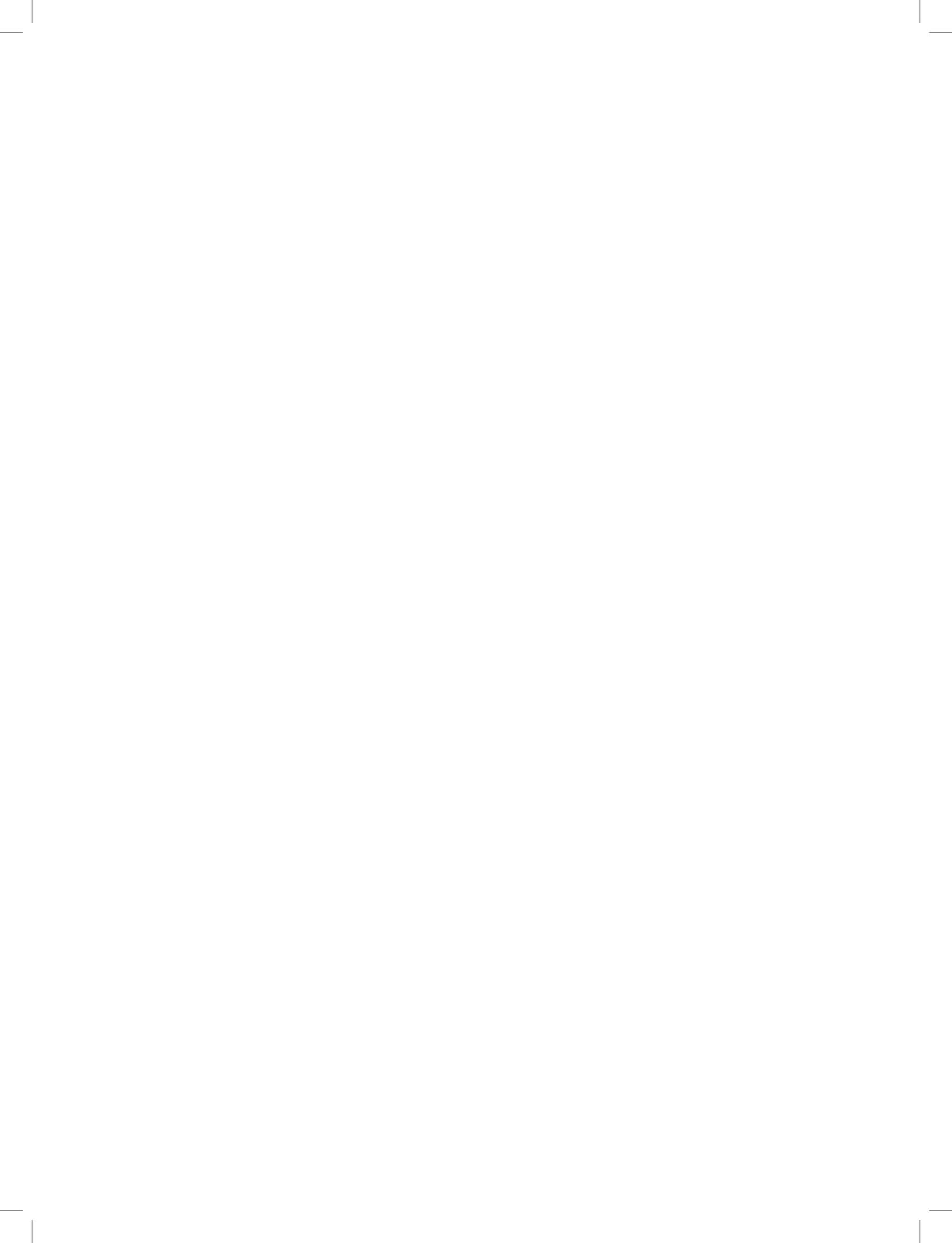
69-50% = Regular, fue difícil el proceso de aprendizaje y lo logré parcialmente.

49% o menos = Insuficiente, no logré el aprendizaje.

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Conceptuales	Contenidos				
	Identificas los estadígrafos de tendencia central.				
	Identificas los estadígrafos de dispersión.				
	Conoces las características de una curva normal.				

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Procedimentales	Contenidos				
	Analizas críticamente la información que se te presenta.				
	Calculas los estadígrafos de tendencia central para un conjunto de datos organizados en clases y frecuencias.				
	Calculas los estadígrafos de dispersión para un conjunto de datos organizados en clase y frecuencias.				
	Discutes la relación empírica entre estadígrafos de tendencia central y estadígrafos de dispersión.				

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Actitudinales	Contenidos				
	Valoras el trabajo en equipo como elemento que aporta y contrapone ideas en la resolución de problemas.				
	Cumples con las indicaciones dadas para el buen desarrollo de las actividades.				
	Buscas y sugieres soluciones a los problemas planteados.				
	Tienes una actitud positiva hacia el trabajo desarrollado en el bloque.				



BLOQUE IV

Analizas la teoría de conjuntos
y sus aplicaciones



¿Qué aprenderás y cómo organizarás tu estudio?

Bloque IV

10
HORAS

Objetos de aprendizaje que se abordan

Conceptos básicos.
Conjunto
Subconjunto
Elemento
Relaciones y operaciones entre conjuntos.
Teoría de probabilidad.

Desempeños esperados al concluir el bloque

- Identificas los elementos de un conjunto y sus operaciones.
- Analizas y reconoces las operaciones de un conjunto como base para el cálculo de la probabilidad.
- Comprendes las características de marco muestral y un evento como elementos básicos del cálculo de la probabilidad simple.

Competencias disciplinares del campo de las matemáticas

- Expresas ideas y conceptos mediante representaciones lingüísticas y matemáticas relacionadas con las operaciones básicas de la teoría de conjunto y la probabilidad y sus aplicaciones.
- Sigues instrucciones y procedimientos de manera reflexiva y ordenas información de acuerdo con categorías, jerarquías, relaciones con base en los conceptos básicos de la teoría de conjuntos y la probabilidad.
- Sintetizas evidencias obtenidas mediante la experimentación para producir conclusiones y formular nuevas preguntas relacionada con la teoría de conjuntos y la probabilidad.
- Estructuras ideas y argumentos de manera clara, coherente y sintetizas los resultados de la teoría de conjuntos y la probabilidad.
- Defines metas y das seguimiento a tus procesos de construcción de conocimiento sobre la teoría de conjunto y la probabilidad.
- Propones maneras de solucionar un problema o desarrollar un proyecto en equipo, definiendo un curso de acción con pasos específicos.
- Asumes una actitud constructiva, congruente con los conocimientos y habilidades con los que cuentas dentro de distintos equipos de trabajo.

Productos de aprendizaje

En este bloque realizarás los siguientes ejercicios:

- Representación y análisis de conjuntos en diagramas de Venn-Euler.
- Análisis de las relaciones de conjunto de tu cuadro de frecuencias.
- Cálculo de probabilidades de diferentes eventos.
- Determinación de las características de una muestra aleatoria simple.

Introducción

En los bloques anteriores trabajaste un conjunto de datos de tu comunidad. Como ya sabes, los resultados obtenidos únicamente son válidos para quienes entrevistaste (la muestra), no para toda tu comunidad (la población). Si quisieras que estos resultados fueran válidos para toda tu comunidad, tu muestra debió haber sido aleatoria.

En este bloque desarrollarás las bases para hacer un buen muestreo que te permita reflejar, en un conjunto pequeño de datos de las personas entrevistadas, las características y comportamientos de toda tu comunidad (población).

Mapa de objetos de aprendizaje



Como primer paso, vamos a trabajar un poco con la teoría de conjuntos para que puedas ubicar la muestra como un subconjunto de la población, e identificar las características que tienen la muestra y la población.



Actividad de aprendizaje 1

Observa detenidamente el siguiente diagrama de Venn-Euler y responde las preguntas:



1. ¿Cuál es el ser vivo que nada y que vuela y además es un ave?

2. ¿Qué diferencias y similitudes tienen el loro, el pingüino y el pez volador?

3. ¿Qué diferencias y similitudes tienen la avestruz, la mariposa y la ballena? Anota tus conclusiones y discute con tus compañeros.

4. ¿El diagrama te ayudó a contestar las preguntas anteriores? Argumenta por qué.



Sabías que...

Los diagramas de Venn-Euler toman su nombre de los trabajos del inglés John Venn y el suizo Leonhard Euler. Estos diagramas te permiten ver conjuntos y subconjuntos de forma sencilla para realizar operaciones con ellos.

Se componen de un rectángulo que indica el conjunto universal y se denota por 'U'. Dentro de ese rectángulo se pueden insertar tantos conjuntos o elementos (expresados con círculos) como se desee.

El diagrama anterior nos muestra una colección bien definida de objetos o elementos, que pueden ser cualquier cosa: números, personas, letras.

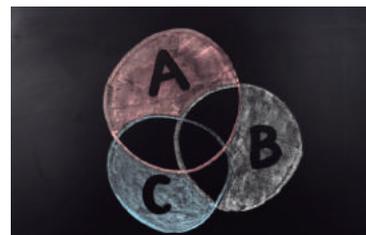
Un conjunto está bien definido si se puede saber con precisión si cualquiera de sus elementos pertenece o no pertenece al conjunto.

En nuestro ejemplo podríamos decir que:

A es el conjunto de las aves.

B es el conjunto de los seres vivos que vuelan.

C es el conjunto de los seres vivos que nadan.



Sabías que...

Los conjuntos se denotan habitualmente con letras mayúsculas. Los objetos que componen el conjunto se llaman elementos o subconjuntos. Se dice que “pertenecen” al conjunto y se denotan en letras minúsculas. La expresión $a \in A$ se lee como “a es elemento de A”, “a está en A”, “a pertenece a A”, “A contiene a a”. Para indicar la noción contraria, que a no es elemento de A se usa el símbolo \notin .



Actividad de aprendizaje 2

Retomando el mismo ejemplo, escribe en lenguaje matemático y coloquial las siguientes afirmaciones:

Ballena \in C: _____

Loro \in A y B: _____

Avestruz no pertenece a B y C: _____

Pez volador no es elemento de A: _____



Aprende más

Relaciones y operaciones entre conjuntos

Para definir correctamente un conjunto hay que considerar las siguientes reglas:

- El conjunto ha de estar bien definido, no debe dar cabida a confusiones como, por ejemplo, que un elemento esté y no esté en el conjunto. Debe ser preciso, por ejemplo: sea B el conjunto de los días de la semana, $B = \{\text{domingo, lunes, martes, miércoles, jueves, viernes, sábado}\}$.
- El orden de los elementos es irrelevante. Pueden aparecer en cualquier orden, por ejemplo: sea M el conjunto de las vocales, $M = \{a, i, o, u, e\}$.



Actividad de aprendizaje 3

Recordarás el ejercicio de las edades de 50 habitantes de tu comunidad que hemos venido desarrollando a partir del bloque II. Es evidente que, además de la edad, esos 50 habitantes cuentan con otras características como son: sexo, si asisten a la escuela, si hablan alguna lengua indígena, entre otras. Supón que los datos de estas características son los que se muestran en la figura.

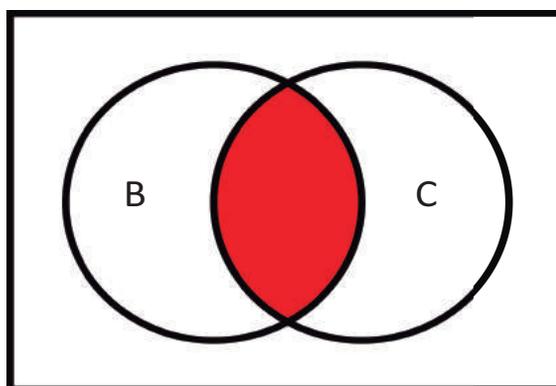


Ahora supón que en las entrevistas que realizaste encontraste que: de los 20 hombres, sólo 10 asisten a la escuela ¿cómo representarías esta información en un diagrama de Venn-Euler?



Si, además, quisieras representar que de los 10 hombres que asisten a la escuela 2 de ellos hablan lengua indígena ¿cómo incorporarías esta información en el diagrama anterior?

Si te fijas, en la primera representación que hiciste tu conjunto B (hombres) y tu conjunto C (asisten a la escuela) comparten elementos en común, pues hay 10 hombres que asisten a la escuela. A esta relación entre subconjuntos se le llama *intersección* y se denota por: $B \cap C$, que se lee: B intersección C. O el conjunto de elementos que están en B y en C.



Bloque IV

Analizas la teoría de conjuntos y sus aplicaciones

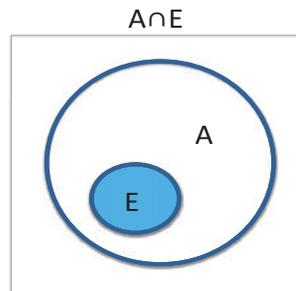
Sin embargo, en nuestro ejemplo establecimos dos intersecciones que son: hombres (conjunto B) que asisten a la escuela (conjunto C) y que hablan lengua indígena (conjunto D). Esto se denota: $B \cap C \cap D$.

Realiza el diagrama que represente a los hombres que asisten a la escuela y hablan lengua indígena.



Hay otras formas de intersección, dependiendo de los conjuntos que te interesa formar. Por ejemplo, las habitantes de tu comunidad que son madres, estarán siempre dentro del conjunto de mujeres, lo cual también denota una intersección.

Si A es el conjunto de mujeres y E es el conjunto de madres. E denota a las mujeres que son madres.



Actividad de aprendizaje 4

Piensa ahora ¿qué tienen en común los diagramas anteriores?, ¿en qué son diferentes? Argumenta tu respuesta.



Sabías que...

Se pueden intersectar tantos conjuntos como sea necesario. No hay que olvidar que es importante definir los elementos que representa cada conjunto y los elementos que comparte con los demás. Recuerda además que no debes excederte de la cantidad de elementos que están definidos para el conjunto.

Así como la suma, la resta o la multiplicación son relaciones entre números, la intersección es una relación entre conjuntos. En el desarrollo de las siguientes actividades te mostraremos cómo utilizarla.



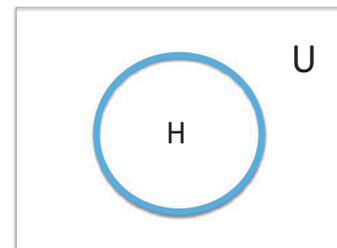
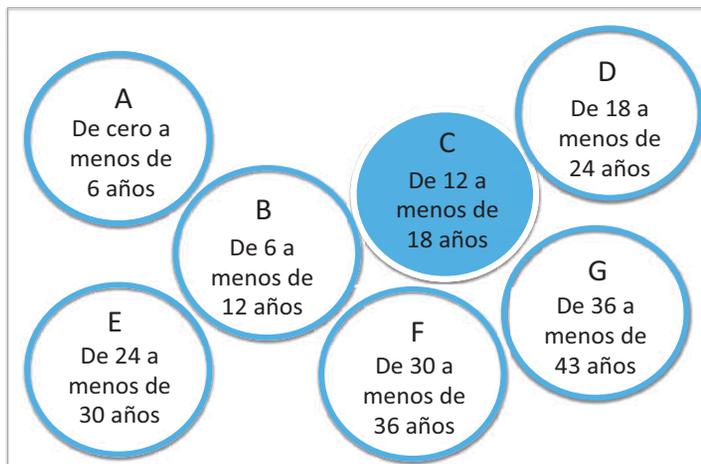
Actividad de aprendizaje 5

Observa el diagrama de las edades de los habitantes de tu comunidad y responde las siguientes preguntas:

1. ¿Qué representa el círculo sombreado dentro del primer diagrama?

2. ¿Qué representa el círculo del segundo diagrama?

3. Indica el número de habitantes que tienen más de 44 años.

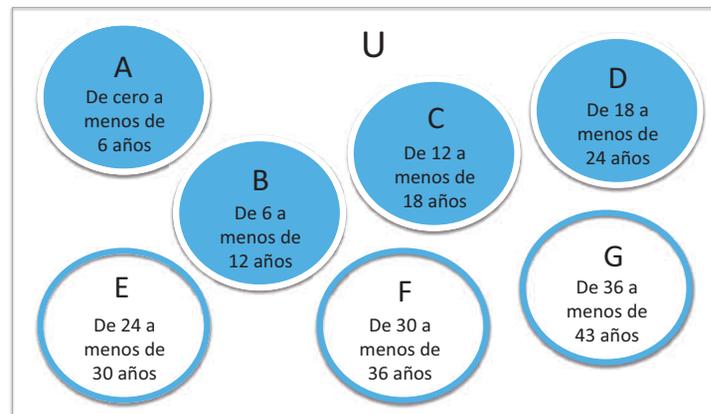


Conjunto de personas mayores de 44 años

Como pudiste observar, en el primer diagrama no existe ninguna persona de más de 44 años. Sería un conjunto de cero elementos que se conoce como conjunto vacío y se denota como $H:0$.

Veamos ahora otras relaciones entre conjuntos: *la unión*.

Si quisiéramos saber la población que está en edad típica de cursar algún nivel educativo, podríamos sombrear los conjuntos A, B, C y D. Esto significa que la población en edad ideal de cursar algún nivel educativo es la unión del conjunto A, B, C y D y se denota como: $A \cup B \cup C \cup D$ y se lee como A unión B unión C unión D, el conjunto de elementos que están en A o en B o en C o en D.



Como seguramente ya te diste cuenta, mientras que en la intersección la letra clave es “y”, en el caso de la unión, lo es la letra “o”.



Actividad de aprendizaje 6

Menciona tres casos de conjuntos que tengan elementos en unión (o) y en intersección (y).

Unión	Intersección



Sabías que...

Los conjuntos que hemos visto hasta ahora son conjuntos finitos porque se puede determinar su “cardinalidad”. Sin embargo, esto no es posible con todos los conjuntos. A los que no se les puede definir su último elemento se les llama conjuntos infinitos, ejemplo: números naturales, enteros positivos y negativos, números fraccionarios, decimales, racionales e irracionales.

Al conjunto que tiene un solo elemento se le conoce como conjunto unitario por ejemplo el cero.



Cardinalidad: la cardinalidad de un conjunto es el número de elementos que contiene. Para conocerla debe ser posible enumerar cada elemento.

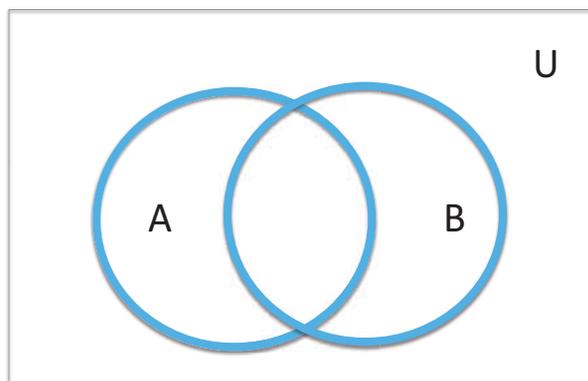


Actividad de aprendizaje 7

El siguiente diagrama representa el universo de hombres de tu comunidad, el subconjunto A comprende a los que están casados y el subconjunto B a los que tienen hijos.

Colorea:

1. Los hombres que están **únicamente** casados (de color amarillo).
2. Los hombres que están casados **y** tienen hijos (de color gris).
3. Los hombres que **solamente** tienen hijos (de color azul).



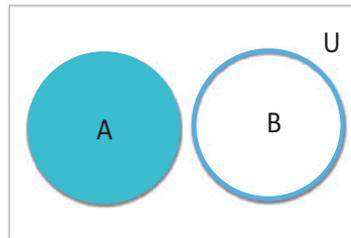


Actividad de aprendizaje 8

A la diferencia entre en el conjunto de los hombres casados y el conjunto de hombres que tienen hijos se le denomina **diferencia de conjuntos**; y se denota por: $A - B$. Se lee: A diferencia B o, A menos B y representa los elementos que pertenecen al conjunto A, pero no pertenecen al conjunto B.

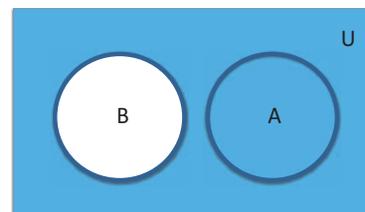
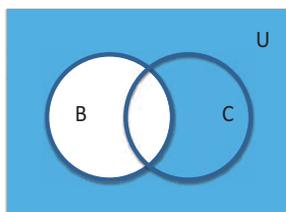
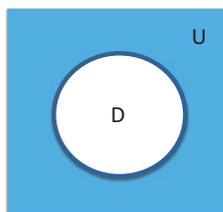
Veamos ahora otro conjunto: la **diferencia simétrica** en este conjunto representan al mismo tiempo los elementos que están en A o en B, pero no en ambos, en nuestro ejemplo sería la parte de rojo y la parte de azul al mismo tiempo.

Por último, cuando los conjuntos A y B no tienen elementos en común, la diferencia $A - B$ en el diagrama sería:



Retomemos nuestra primera actividad de aprendizaje de este bloque, donde clasificamos en el conjunto A a las mujeres, en B a los hombres, en C a los que asisten a la escuela y en D a los que hablan alguna lengua indígena.

Analiza detalladamente los siguientes diagramas y explica qué representan las partes sombreadas de acuerdo con la información del párrafo anterior.



Seguramente notaste que ahora en nuestro universo lo sombreado no corresponde a los conjuntos establecidos, sino a los elementos que no pertenecen a alguno de estos conjuntos. Por ejemplo, a los elementos fuera del conjunto D se les llama complemento de D y se denota como: D^c



Actividad de aprendizaje 9

Para concluir esta sección te pedimos que analices más a fondo tu cuadro de frecuencias utilizando las siguientes relaciones: unión, interacción, diferencia y complemento.

Recuerda que no se trata de establecer relaciones y luego pensar. Por el contrario, primero debes pensar qué buscas y para qué, y luego establecer las relaciones que te ayuden a esto. Sabemos que probablemente estés pensando: “ya estoy harto de Bachelard, ¿cómo puedo saber qué relaciones establecer, si no sé qué estoy buscando?”.

Bueno, pues tienes razón; para aclarar este punto te daremos dos pistas:

Los elementos que integran un conjunto deberán permitirte representar las características que debe reunir tu muestra para poder generalizar las observaciones de esta muestra (los 50 entrevistados) a toda la población (todos los habitantes de tu comunidad).

Debes tener una idea de las características que quieres analizar de la población de tu comunidad, para sobre esa base, diseñar una muestra con esas mismas características. Ejemplos:

- a. Si estimas que en tu comunidad hay 10% de hombres que están en edad de estudiar y hablan lengua indígena, deberás hacer una muestra que contemple ese mismo porcentaje.
- b. Si supones que en tu comunidad hay un porcentaje de mujeres que son madres, deberás buscar que tu muestra cuente con esa misma proporción de mujeres que son madres.

Como habías podido observar, en los incisos a) y b) dijimos estimar y suponer y nunca afirmamos que conocías o sabías. En otras palabras, de lo que se trata es de:

- Tenemos una idea del perfil del universo que quieres estudiar (las características de la población que quieres estudiar) y;
- Buscar que la muestra que elijas refleje fielmente esas características.

A la actividad del punto 1 se le llama *definir el marco muestral*. En el ejercicio que hemos venido desarrollando, algunas de las características de la población de tu comunidad las podrías obtener del INEGI.

Respecto al punto 2, lo que tendrás que hacer, una vez identificado el marco muestral, será garantizar que todas las personas de su comunidad tengan la misma probabilidad de ser elegidas en la muestra.



Sabías que...

El tipo más sencillo de marco muestral es una lista de elementos de la población con la información de contacto adecuada. Por ejemplo, la lista con los números que asignaste a cada habitante de tu comunidad.



Aprende más

Teoría de la probabilidad

Dicho todo esto, ¡pon manos a la obra! En el último párrafo usamos, por primera vez, la palabra: probabilidad.

¿Qué es la probabilidad? Para tener una primera aproximación a este concepto, juguemos un rato:



Actividad de aprendizaje 10

Seguro que alguna vez has lanzado volados con tus amigos.

¿Cuál es la probabilidad de obtener un sol en el primer volado?

1. Anota tu respuesta _____

2. Y, ¿cuál es la probabilidad de obtener sol en los siguientes 9 volados?

3. Ahora lanza 10 volados y anota los resultados:



Como te habrás dado cuenta, en un volado sólo hay dos posibilidades: águila o sol. Lo que significa que ambos resultados son igualmente posibles. Si es así, la probabilidad de cada evento (águila o sol) es $1/n$ donde $n = 2$, águila o sol. Esto se llama *probabilidad simple*.



Sabías que...

La probabilidad de un evento simple se denota como $P(A)$, siendo A el evento simple en cuestión.



Actividad de aprendizaje 11

¿Te pareció muy fácil el ejercicio anterior?

¿Qué sucedería si en vez de dos resultados (águila o sol) tuvieras 87 canicas en una bolsa y 68 de ellas fueran verdes y el resto rojas?. Si escoges una, ¿cuál es la probabilidad de que esta canica sea verde? Escribe tu respuesta y discútela con tus compañeros y asesor.



Seguramente pensaste en hacer la misma operación que en el caso de los volados y te diste cuenta que el número de canicas verdes dentro de la bolsa no es el mismo que el de canicas rojas. Entonces, ¿cómo calcularías la probabilidad cuando los eventos no son equiprobables, es decir, cuando NO tienen la misma probabilidad?

¡Pues muy fácil! cuando los eventos no son equiprobables, la probabilidad de un evento A está dada por el número de casos posibles (número de canicas verdes) entre el número de casos totales (total de canicas).

$P(A) = \# \text{ de casos posibles} / \# \text{ total de casos}$. A este tipo de probabilidad se le conoce como *probabilidad clásica*.



Equiprobable: sucesos de un experimento que tienen la misma probabilidad de ocurrencia.

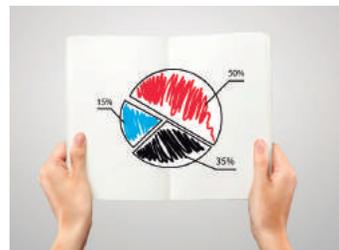


Actividad de aprendizaje 12

Ahora que ya sabes cómo calcular probabilidades y retomando tu ejercicio de los 50 habitantes de tu comunidad, considera la siguiente información:

Se definen los siguientes eventos:

- A: Es mujer
- B: Es hombre
- C: Es un habitante que asiste a la escuela
- D: Es un habitante que habla lengua indígena.



Responde lo siguiente:

1. ¿Cuál es la probabilidad de elegir un habitante de tu comunidad que sea mujer? _____
2. ¿Cuál es la $P(B)$, $P(C)$ y $P(D)$? _____
3. ¿Es igual la probabilidad de elegir una mujer que un hombre en tu comunidad? Explica por qué: _____



Actividad de aprendizaje 13

Reflexiona el siguiente problema:

Dentro de dos semanas tendrás una entrevista con el presidente municipal en la que se te solicitará un informe de las edades de los habitantes de toda tu comunidad. Para asistir a esta reunión has preparado tres estrategias:

1. Generalizar el informe estadístico de tu muestra al resto de la población de tu comunidad.
2. Obtener todas las edades de la población de tu comunidad entrevistando a todas y todos sus habitantes.
3. Obtener la información de tu comunidad de los censos de población del INEGI.

¿Qué ventajas y desventajas tiene cada una de estas estrategias? Explica ampliamente.

Seguramente todos tus argumentos son válidos, sin embargo, cada una de estas estrategias tiene sus limitaciones:

La primera: porque no estás seguro de que se pueda generalizar el ejercicio estadístico de tus entrevistados al resto de la población de tu comunidad en la medida en que los datos que obtuviste en tu primera encuesta en campo fueron de las primeras 50 personas que te encontraste y no de una muestra aleatoria.

La segunda: porque consideras que es imposible llevarlo a cabo desde un punto de vista práctico, pues los recursos humanos y financieros con que cuentas no serían suficientes.

La tercera: porque la información que puedes obtener del censo no está actualizada.

Así que por todas estas dudas recurres a la asesoría de tu maestro de Probabilidad y Estadística para que te oriente en tomar una decisión estadística para este informe.



Actividad de aprendizaje 14

Para que tu aprendizaje sea significativo, tu maestro te recomienda que observes la población de tu comunidad y contestes las siguientes preguntas:

1. ¿Qué puedes describir de sus habitantes? Explica ampliamente al menos 5 características:

2. ¿Cómo se te ocurre que puedes conservar dichas características en un conjunto más pequeño de la población que observas? Discute con tus compañeros

3. Si consideras que tu comunidad cuenta con 500 habitantes repartidos a lo largo del territorio. ¿Cómo podrías seleccionar a aquellos que en una muestra pudieran representar a toda tu comunidad?

Probablemente te diste cuenta de que el subconjunto de 50 personas que entrevistaste debió tener las características de la población de toda tu comunidad. A este subconjunto se le podría llamar muestra aleatoria. Una muestra cuyos resultados serán representativos de la población general.

Para que esa muestra sea representativa de la población debes preguntarte si es confiable la técnica de muestreo que escogiste.

Existen diferentes técnicas de muestreo.



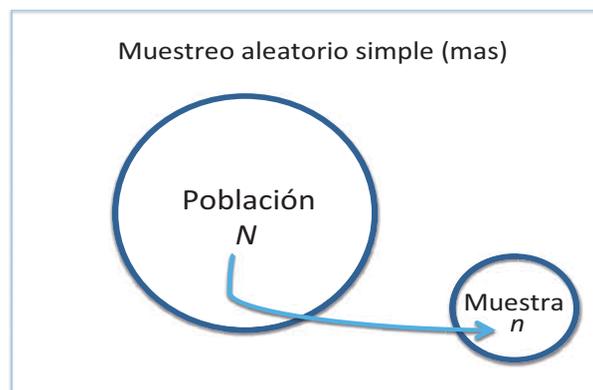
Actividad de aprendizaje 15



Responde lo siguiente:

Si tú les dieras un número a cada uno de los habitantes del municipio y esos números los anotaras en una papeleta. Luego introduciras todos esos números en una bolsa y los revolveras muy bien, ¿qué probabilidad tendrían los habitantes de toda tu comunidad de ser escogidos para que los entrevistaras? Justifica ampliamente tu respuesta.

A la técnica de muestreo en la que los elementos tienen la misma probabilidad de ser seleccionados se le conoce como *muestreo aleatorio simple*. Este tipo de muestreo se fundamenta en el siguiente esquema:



Selección aleatoria de los elementos muestrales con probabilidades de selección en cualquier extracción iguales y sin reemplazo.



Sabías que...

El muestreo aleatorio simple es una técnica de muestreo probabilístico en la que todos los sujetos tienen una probabilidad conocida, distinta de cero, de ser seleccionados, es una técnica que utiliza métodos aleatorios para la selección.

Un ejemplo de este tipo de muestreo son los volados. Cuando echaste los volados, probablemente no te salieron 5 águilas y 5 soles. Sin embargo, si echas más y más volados, seguramente las águilas y soles tenderían a ser 50 y 50%, aunque no exactamente en esta proporción.



A esa diferencia que puede existir respecto al 50 - 50% se le llama error aleatorio o de muestreo. Como su nombre lo indica, por azar, que no implica que la muestra no sea representativa.

Cierre de bloque IV

En este bloque trabajaste la teoría de conjuntos y probabilidad para la determinación de la muestra adecuada para tu población.

Este proceso te llevó a manejar los conjuntos y sus relaciones, es decir, la unión, intersección, diferencia y complemento que puedes encontrar en cualquier población.

Asimismo conociste una noción de probabilidad, cómo aplicarla y la importancia de conservar las proporciones dentro de la muestra.

Autoevaluación

Lee detenidamente las preguntas y responde colocando una X en el nivel de avance que consideras que has logrado a lo largo del bloque IV.

Interpretación del nivel de avance:

100-90% = Excelente, logré el aprendizaje de manera independiente.

89-70% = Bueno, requerí apoyo para construir mi aprendizaje.

69-50% = Regular, fue difícil el proceso de aprendizaje y lo logré parcialmente.

49% o menos = Insuficiente, no logré el aprendizaje.

Bloque IV

Analizas la teoría de conjuntos y sus aplicaciones

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Conceptuales	Contenidos				
	Identificas las relaciones de un conjunto de datos (unión, interacción, diferencia y complemento).				
	Identificas las características de una muestra.				
	Elaboras ejemplos de conjuntos y sus operaciones.				

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Procedimentales	Contenidos				
	Defines un marco muestral para un conjunto de datos.				
	Comprendes el término de probabilidad y cuál es su utilidad en la selección de una muestra.				
	Organizas información de acuerdo con los conceptos básicos de la teoría de conjuntos.				

		Nivel de avance			
		100-90%	89-70%	69-50%	49% o menos
Actitudinales	Contenidos				
	Valoras el trabajo en equipo como elemento que aporta y contrapone ideas en la resolución de problemas.				
	Cumples con las indicaciones dadas para el buen desarrollo de las actividades.				
	Buscas y sugieres soluciones a los problemas planteados.				
	Tienes una actitud positiva hacia el trabajo desarrollado en el bloque.				

Cardinalidad: la cardinalidad de un conjunto es el número de elementos que contiene. Para conocerla debe ser posible enumerar cada elemento.

Censo: el estudio de todos los elementos de una población. En México se realiza el censo de población cada 10 años.

Clase: intervalos contruidos para agrupar la información.

Clase mediana: clase a la que pertenece el valor de la mediana.

Clase modal: en una tabla de datos agrupados, es la clase o intervalo que tiene la mayor frecuencia de observaciones.

Equiprobable: sucesos de un experimento que tienen la misma probabilidad de ocurrencia.

Estadística: es la ciencia que te ayuda a recopilar, analizar e interpretar información muy diversa. Busca darle sentido a ésta información para ayudar a tomar decisiones.

Estadística descriptiva: conjunto de procedimientos para organizar, resumir y analizar un conjunto de datos.

Histograma: representación gráfica de una distribución de frecuencias de una variable continua. Describe el comportamiento de un conjunto de datos en cuanto a su tendencia central, forma y dispersión. Está formado por un conjunto de rectángulos unidos, cuya base es igual a la amplitud del intervalo de clase y a la longitud proporcional a la frecuencia.

Inferencia estadística: conjunto de procedimientos cuya finalidad es obtener conclusiones respecto a la población a partir de datos observados en muestras.

Muestra: subconjunto representativo de la población a partir del cual se pretende realizar inferencias respecto a la población de donde procede.

Población: conjunto de datos o elementos que interesa analizar.

Polígono de frecuencias: gráfico utilizado para representar una distribución de frecuencias de una variable numérica, teniendo en cuenta su marca de clase.

Tabla de distribución de frecuencias: tabla donde se agrupan los datos en categorías mutuamente excluyentes que indican el número de observaciones en cada categoría.

Variabilidad: se refiere a los cambios que presenta una variable.

Variable: característica de la población o de la muestra cuya medida puede cambiar de valor.

¿Con qué conocimientos cuento?

Problema 1

Datos:

2 gotas por cada 10 kg.

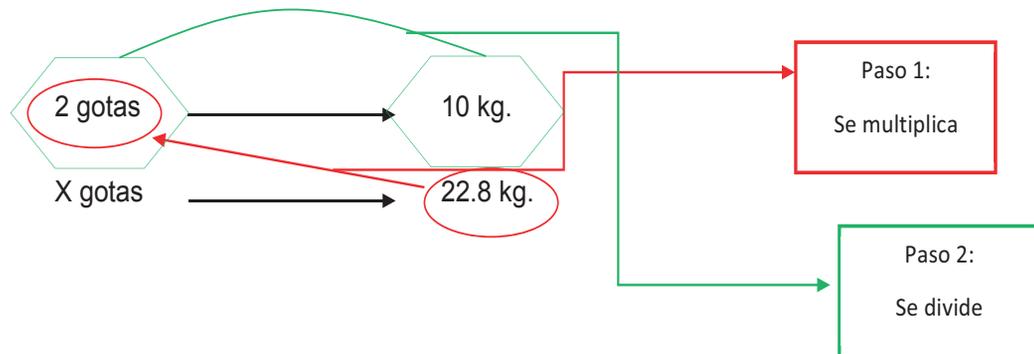
peso del niño: 22.8 kg

Problema: número de gotas deben administrarse

Planteamiento: regla de tres

... es a ...
2 gotas ----- 10 kg.
X gotas ----- 22.8 kg.

Solución:



Paso 1: se multiplica el valor de los kilogramos por el número de gotas establecidas para cada kilogramo $22.8 \text{ kg} * 2 \text{ gotas} = 45.6 \text{ kg} * \text{gota}$.

Es importante destacar que las unidades de las medidas en este caso kg y gotas también se multiplican y forman: "kg*gota".

Paso 2: se divide el resultado de la multiplicación del Paso 1 entre el número de kilogramos establecidos para las 2 gotas. $(45.6 \text{ kg} * \text{gota}) / (10 \text{ kg}) = 4.56 \text{ gotas}$.

Es importante destacar que en una división, las unidades de las medidas en este caso kg y gotas también se dividen, como sabes si en el numerador y en el denominador aparecen las mismas unidades, éstas se eliminan por eso en este paso únicamente se conserva como unidades: "gotas"

Resultado:

La mamá del niño con peso de 22.8 kg. deberá administrarle 4.56 gotas.

Sin embargo, no hay fracciones de gotas en las medicinas por lo que deberá de administrarle: 5 gotas completas.

Problema 2

Datos:

12 botes de pintura azul de $\frac{1}{2}$ lt. cada uno.

Barda: 90 m. largo X 80 cm. alto.

Edificio: 15 m alto 40 m largo.

botes de 2 lts. de pintura

Problema: número de botes de 2lts. para pintar el edificio

Planteamiento 1:

0. Convertir cm. a m. para unificar unidades de medida

1. Sacar el área de la barda

2. Sacar el área de la pared del edificio

3. Hacer una regla de tres

4. Obtener el total de los lt. de pintura necesarios para pintar la pared

5. Dividir el total de los lt del paso anterior entre 2lt que es el tamaño de los botes de pintura que se tienen

6. Obtener el número de botes de pintura que se necesitan.

Planteamiento 2:

0. Convertir cm. a m. para unificar unidades de medida

1. Establecer los datos para realizar la regla de tres compuesta e inversamente proporcional

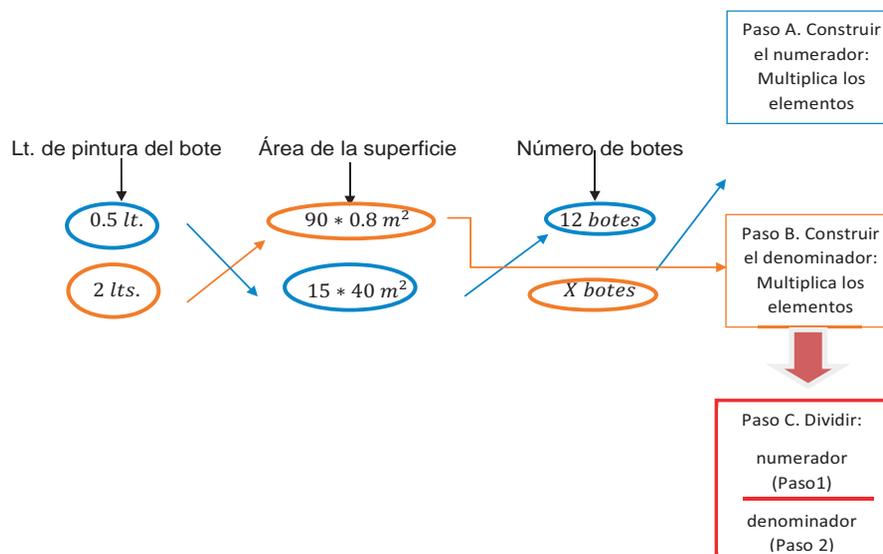
2. Resolver

3. Obtener el número de botes de pintura que se necesitan.

Solución para el planteamiento 2:

Paso 0: convertir cm en metros ($80 \text{ cm} = 80 \cdot 0.01 = 0.8 \text{ m}$)

Paso 1: establecer los datos para realizar la regla de tres compuesta e inversamente proporcional.



Paso 2: resolver

$$\text{Paso A: } (0.5 \text{ lt.}) * (15 * 40 \text{ m}^2) * (12 \text{ botes}) = 3600 \text{ lts.} * \text{m}^2 * \text{bote}$$

$$\text{Paso B: } (2 \text{ lts.}) * (90 * 0.8 \text{ m}^2) = 144 \text{ lts.} * \text{m}^2$$

$$\text{Paso C: } \frac{3600 \text{ lts.} * \text{m}^2 * \text{botes}}{144 \text{ lts.} * \text{m}^2} = 25 \text{ botes}$$

Paso 3: resultado

25 botes de dos litros se necesitan para pintar la pared que no tiene ventanas del edificio.

Problema 3

20 de los 30 compañeros del salón se fueron de viaje. ¿Qué proporción de alumnos salieron de viaje?

Datos:

30 compañeros en el salón

20 compañeros se fueron de viaje

Problema: proporción de compañeros que salieron de viaje

Planteamiento: la proporción indica que del total, una parte tiene una característica en particular. En el caso de los alumnos en el salón de clases, son 30 en total. Y sólo 20 de ellos se fueron de viaje. Por tanto la proporción es 20 a 30 alumnos de viaje.

$$\text{Solución: } \frac{\text{Casos posibles}}{\text{Casos totales}} = \frac{\text{Alumnos de viaje}}{\text{Alumnos del salón}} = \frac{20}{30} = \frac{2}{3}$$

Resultado: $\frac{2}{3}$ de los alumnos se fueron de viaje.

Problema 4

Datos:

Dos artículos

Precio de venta unitario: 75 pesos

Ganancia 10%

Problema: costo final unitario

Planteamiento:

1. Calcular el 10% de ganancia sobre un artículo, sabiendo que al dividir por 1.10, obtenemos el precio unitario (1 entero) más la ganancia del 10% (0.10 enteros)
2. A la división anterior, multiplicarla por dos.

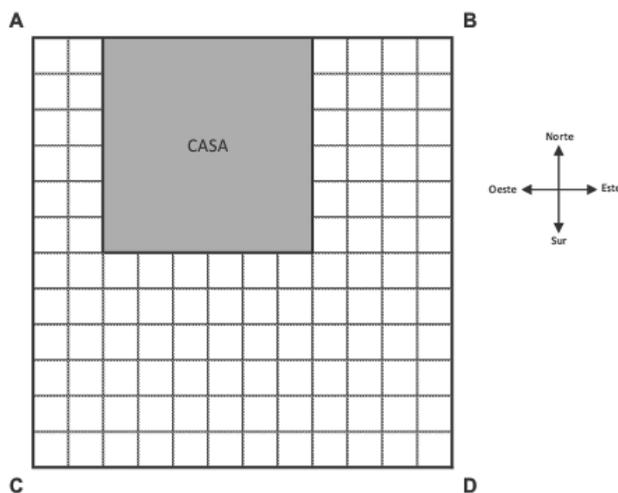
Solución:
$$\frac{75 \text{ precio unitario}}{1.10} = 68.18 \text{ precio final unitario}$$

68.18 precio final unitario x 2 = 136.36 pesos

Resultado: el precio final unitario es decir por cada artículo es de 68.18 pesos, mientras que el precio final por ambos artículos es de 136.36 pesos.

Problema 5

a)



b) Puesto que no se plantan árboles en los segmentos donde está la cerca, ni se plantan árboles en el área de la casa de la tía Rosario, entonces tenemos 4 líneas de 11 árboles cada una (en total son 44 árboles) y 7 líneas de 5 árboles cada una (en total son 30 árboles) por lo tanto 44 + 30 árboles es igual a necesitar: 74 árboles para reforestar el terreno de tía Rosario.

Problema 6

c) $9 / 5 \text{ } ^\circ\text{C} - \text{ } ^\circ\text{F} + 32 = 0$

Problema 7

- a) 0 (CERO)
- b) 0 (CERO)
- c) 1 / 7
- d) 0 (CERO)

Actividad de aprendizaje 1

Los temas sobre los cuales te interesaría reunir información pueden ser muy diversos. Te proponemos temas como la población con estudios universitarios (educación); población con enfermedades crónico degenerativas (salud); población con trabajo formal (trabajo); actividades económicas (economía) entretenimiento (cultura).

Fuentes directas: entrevistas y encuestas. Fuentes secundarias (nacionales e internacionales): el Instituto Nacional de Estadística y Geografía (INEGI), el Consejo Nacional de Población (CONAPO), Organización para la Cooperación y Desarrollo Económicos (OCDE) y la Organización Mundial de la Salud (OMS), La Organización de las Naciones Unidas para la Educación, la Ciencia y la UNESCO, la Agricultura (FAO).

Actividad de aprendizaje 2

Si tuvieses que estudiar un tema educativo las respuestas serían:

Delimitar población: Estudiantes universitarios. De 18 a 20 años. Que estudian ingeniería.

Delimitación de la región geográfica en donde se estudiará la población. Siguiendo el ejemplo sobre educación, estudiantes universitarios del Estado de México. De 18 a 20 años. Que estudian ingeniería.

Definir muestras lo suficientemente representativas para evitar tener que encuestar o entrevistar a toda la población, pues no siempre se cuentan con el tiempo o los recursos suficientes para estudiar a todos los miembros de una población.

Actividad de aprendizaje 3

Para poder establecer si es válido o no para todos los demás Telebachilleratos del Estado, resulta imprescindible saber que tan representativos en términos numéricos son con respecto al total de Telebachilleratos existentes.

Si estos Telebachilleratos, son representativos no solo en número sino también en características de las condiciones del resto de Telebachilleratos.

Actividad de aprendizaje 4

Guarda tu listado porque más adelante te podrá ser de utilidad.

Actividad de aprendizaje 5

La OMS obtuvo esta información de los datos del Fondo de las Naciones Unidas para la Infancia (UNICEF), la Encuesta Nacional de Salud y Nutrición 2012 y datos que la misma OMS recoge y analiza tomando los registros administrativos de las Secretarías y Ministerios de Salud de todos los países para poder realizar análisis de la situación de obesidad.

Para poder hacer esta afirmación la OMS obtuvo una muestra lo suficientemente representativa que le permitió hacer una encuesta y que los resultados de la misma fueran posibles de inferir o generalizar para la población total.

Actividad de aprendizaje 6

1. Horas promedio de estudio y calificación promedio, pues dependiendo de las horas promedio que las personas destinen para estudiar así será el desempeño académico de las mismas. Asimismo la calificación promedio permite establecer el nivel del desempeño de cada estudiante.
2. No, el promedio de calificaciones no puede depender de la estatura, pues no existe una relación directa entre cuanto mide una persona y su promedio de calificaciones.
3. El promedio de calificaciones han ido tendiendo a bajar a lo largo de los seis semestres, el número de alumnos se incrementó al igual que la edad promedio.

Actividad de aprendizaje 7

1. Variables discretas: Número de alumnos, número de alumnos casados, número de alumnos que son padres o madres, número de estudiantes que trabajan. Variables continuas: Promedio de calificaciones, edad promedio, horas promedio de estudio, estatura promedio de los varones y estatura promedio de las mujeres.
2. a. Temperaturas registradas cada hora por el observatorio.
d. Edad de los estudiantes de tu escuela

Actividad de aprendizaje 8

Las calificaciones pueden depender del tiempo que dedican a estudiar los alumnos de telebachillerato. Es menos probable que dependan de la estatura de los estudiantes.

Actividad de aprendizaje 9

- Variables independientes: Sobrepeso y obesidad, inseguridad y corrupción y problemas en el embarazo, fumar antes de los 12 años.
- Variables dependientes: Enfermedades de arterias coronarias y diabetes, el crecimiento de la economía, el bajo peso, cáncer pulmonar.

Actividad de aprendizaje 1

- 1.- Los datos están revueltos
- 2.- Hay pocos datos repetidos (por ejemplo el 26, se repite dos veces)
- 3.- El datos menor es el 2
- 4.- El dato mayor es el 43
- 5.- Los datos no son consecutivos, hay saltos

Actividad de aprendizaje 2

La tabla permite observar ordenados los datos por lo que se ven más claras las cinco características descritas antes. Además, sí se logran distinguir nuevas características que antes no observaba, por ejemplo hay otros números repetidos además del 26 como el 6, 13, 14, 20, 21 y otros. Otra característica que se puede observar más claramente es que el número que más se repite es el 6.

Actividad de aprendizaje 3

Nivel educativo	Número casos
Primaria	8
Secundaria	7
Edad votar	30

El cuadro de conteo nos resume la información en tres grupos, así se puede dar información más concreta en vez de ver cada dato por separado.

- Preguntando a los habitantes de mi comunidad y contaría a todos los que son menores de 17 años. El método que seguiría sería primero preguntar a las personas su edad, y si entra en el programa "Crece más", lo cuento, sino, no.
- Pues sí fue diferente porque seleccioné a ciertas personas y no a cualquiera que me encontrara en la calle. No todos habitantes entran en el programa.

Actividad de aprendizaje 4

Edades de 50 habitantes de tu comunidad				
Edad	Frecuencia absoluta (f_a)	Frecuencia absoluta acumulada (f_{aa})		
2	1	1	f_{aa} en personas de 18 años o más	
3	1	2		
5	1	3		
6	4	7		
8	1	8		
9	1	9		
11	2	11		
12	1	12		
13	4	16		
14	2	18		
16	1	19		
17	1	20		
18	1	21		1
20	2	23		3
21	3	26		6
22	2	28		8
23	3	31		11
24	2	33	13	
26	2	35	15	
27	1	36	16	
28	1	37	17	
29	2	39	19	
30	2	41	21	
31	1	42	22	
32	2	44	24	
33	1	45	25	
35	1	46	26	
36	1	47	27	
37	1	48	28	
39	1	49	29	
43	1	50	30	
Total de frecuencias				

Actividad de aprendizaje 5

Edades de 250 personas de "El Capulín"					
Edad	Frecuencia absoluta (f_a)	Frecuencia absoluta acumulada (f_{aa})	Edad	Frecuencia absoluta (f_a)	Frecuencia absoluta acumulada (f_{aa})
1	1	1	48	5	125
2	1	2	49	2	127
3	3	5	50	6	133
4	3	8	51	3	136
5	5	13	52	3	139
6	2	15	53	4	143
7	5	20	54	1	144
8	5	25	55	2	146
9	3	28	56	2	148
10	3	31	57	1	149
11	5	36	58	1	150
12	2	38	59	1	151
13	3	41	60	3	154
14	2	43	61	3	157
15	1	44	62	2	159
16	2	46	63	5	164
17	5	51	64	2	166
18	1	52	65	3	169
19	5	57	66	1	170
20	3	60	67	2	172
21	4	64	68	3	175
22	1	65	69	3	178
23	1	66	70	2	180
24	1	67	71	3	183
25	1	68	72	2	185
26	2	70	73	3	188
27	3	73	74	2	190
28	4	77	75	2	192
29	1	78	76	7	199
30	2	80	77	3	202
31	2	82	78	3	205
32	3	85	79	3	208
33	2	87	80	3	211
34	2	89	81	5	216
35	2	91	82	1	217
36	2	93	83	4	221
37	6	99	84	7	228
38	3	102	85	2	230
40	1	103	86	5	235
41	1	104	87	2	237
42	1	105	88	6	243
43	4	109	89	4	247
44	7	116	90	3	250
45	3	119			
47	1	120			

1. Las diferencias que puedes observar son las siguientes:

- Una es mucho más grande que la otra, es decir, más datos pues se entrevistaron a más personas.
- Tiene más variedad de número, llegan a 90.
- El número que más se repite es el 84, tiene 7 repeticiones.
- La menor edad es 1 y la mayor es 90.
- No están todos los números consecutivos pero es más difícil notarlo porque los saltos no son evidentes.

Las similitudes que puedes observar son las siguientes:

- Son datos sobre edades de las poblaciones que representan respectivamente.
- Los datos de origen están desordenados.
- Independientemente del número de datos los procedimientos estadísticos se pueden aplicar igual.

2. Sí, pues nos dijeron que ambas muestras fueron extraídas de poblaciones reales. Además, el tamaño está en proporción con el número de habitantes de cada una de las comunidades. Sin embargo, para que una muestra sea buena y representativa debe considerar ciertos criterios de selección que serán revisados en el Bloque 4.

3. Pareciera que sí, entre más grandes mejores las muestras, pero no es así, las muestras deben ser de un tamaño adecuado y deben ser representativas de la población es decir, que la variedad de elementos de una población estén reflejados en la muestra.

Una de las formas de comparar más usada es comparar en porcentajes; de esta manera sería lo mismo cualquier población o muestra de cualquier número de elementos.

Actividad de aprendizaje 6

Edades de 250 personas de “El Capulín”					Edades de 250 personas de “El Capulín”				
Edad	Frecuencia absoluta (f_a)	Frecuencia absoluta acumulada (f_{aa})	Frecuencia relativa	Frecuencia relativa acumulada	Edad	Frecuencia absoluta (f_a)	Frecuencia absoluta acumulada (f_{aa})	Frecuencia relativa	Frecuencia relativa acumulada
1	1	1	0.4	0.4	47	1	120	0.4	49.1
2	1	2	0.4	0.8	48	5	125	2	51.1
3	3	5	1.2	2.0	49	2	127	0.8	51.9
4	3	8	1.2	3.2	50	6	133	2.4	53.3
5	5	13	2	5.2	51	3	136	1.2	54.5
6	2	15	0.8	6	52	3	139	1.2	55.7
7	5	20	2	8	53	4	143	1.6	57.3
8	5	25	2	10	54	1	144	0.4	57.7
9	3	28	1.2	11.2	55	2	146	0.8	58.5
10	3	31	1.2	12.4	56	2	148	0.8	59.3
11	5	36	2	14.4	57	1	149	0.4	59.7
12	2	38	0.8	15.2	58	1	150	0.4	60.1
13	3	41	1.2	16.4	59	1	151	0.4	60.5
14	2	43	0.8	17.2	60	3	154	1.2	61.7
15	1	44	0.4	17.6	61	3	157	1.2	62.9
16	2	46	0.8	18.4	62	2	159	0.8	63.7
17	5	51	2	20.4	63	5	164	2	65.7
18	1	52	0.4	20.8	64	2	166	0.8	66.5
19	5	57	2	22.8	65	3	169	1.2	67.7
20	3	60	1.2	24	66	1	170	0.4	68.1
21	4	64	1.6	25.6	67	2	172	0.8	68.9
22	1	65	0.4	26	68	3	175	1.2	70.1
23	1	66	0.4	26.4	69	3	178	1.2	71.3
24	1	67	0.4	26.8	70	2	180	0.8	72.1
25	1	68	0.4	27.2	71	3	183	1.2	73.3
26	2	70	0.8	28	72	2	185	0.8	74.1
27	3	73	1.2	29.2	73	3	188	1.2	75.3
28	4	77	1.6	30.8	74	2	190	0.8	76.1
29	1	78	0.4	31.2	75	2	192	0.8	76.9
30	2	80	0.8	32.1	76	7	199	2.8	79.7
31	2	82	0.8	32.9	77	3	202	1.2	80.9
32	3	85	1.2	34.1	78	3	205	1.2	82.1
33	2	87	0.8	34.9	79	3	208	1.2	83.3
34	2	89	0.8	35.7	80	3	211	1.2	84.5
35	2	91	0.8	36.5	81	5	216	2	86.5
36	2	93	0.8	37.3	82	1	217	0.4	86.9
37	6	99	2.4	39.7	83	4	221	1.6	88.5
38	3	102	1.2	40.9	84	7	228	2.8	91.3
40	1	103	0.4	41.3	85	2	230	0.8	92.1
41	1	104	0.4	41.7	86	5	235	2	94.1
42	1	105	0.4	42.1	87	2	237	0.8	94.9
43	4	109	1.6	43.7	88	6	243	2.4	97.3
44	7	116	2.8	47.5	89	4	247	1.6	98.9
45	3	119	1.2	48.7	90	3	250	1.2	100

	Porcentaje de habitantes que:	
	Están en edad de votar	Son beneficiarios del programa “60 y + años”
Mi comunidad	60.0%	0%
Comunidad “El Capulín”	79.6%	39.60%

Sí existe similitud, el porcentaje de habitantes que están en edad de votar es mayor del 50% en ambas comunidades. Esto se logró independientemente de la cantidad de entrevistados en cada una de las comunidades, e incluso no afectó que las edades de la comunidad “El Capulín” fueran mayores.

Las ventajas pueden ser las siguientes:

1. Puedes comparar poblaciones que tienen números diferentes de individuos
2. Comparar sólo la parte que es de interés, de esta forma se pueden hacer conjeturas.
3. Es más fácil expresarse en porcentajes cuando se necesita hacer presentaciones o hablar con otras personas, pues a veces es difícil imaginar, millones de personas, pero es más fácil imaginar 50% de personas, uno tiene más idea.

Actividad de aprendizaje 7

1. De la primera tabla se obtiene información sobre obesidad a nivel de países. El porcentaje de la población mayor de 15 años con un índice de masa corporal superior a 30 está en porcentaje por lo que es comparable la información entre países. La fuente de información es la OCDE en los años 1996-2003 que se hizo la investigación. De la segunda tabla se obtiene información del sobrepeso y obesidad de México y el Distrito Federal, los datos se encuentran en porcentajes y en número de personas. Los grupos de análisis son: mujeres mayores de 20 años, hombres mayores de 20 años, niños en edad escolar. La fuente consultada es un trabajo de diversos autores sobre la ENSANUT 2006, con datos del INEGI, 2005.

2. De la primera tabla, la variable es: el porcentaje de población mayor a 15 años con un índice de masa corporal superior a 30. Está en porcentaje. Es continua. De la segunda tabla, la variable es porcentaje de la población con sobrepeso y obesidad y número de personas con sobrepeso y obesidad. La primera es porcentaje la segunda son números. La primera es continua, la segunda es discreta.

3. Hablan sobre temas de salud, en específico sobre sobrepeso y obesidad. Las variables que usan son porcentajes de población. Los elementos de las tablas son los mismos, título, definición de variable, categorías, fuente de información.

5. De la primera, la fuente de información es la OCDE en los años 1996-2003, de la segunda, un trabajo de diversos autores sobre la ENSANUT 2006, con datos del INEGI, 2005.

Actividad de aprendizaje 8

Lo que me está pidiendo el presidente municipal es que agrupe la información;

- Que incluya en el primer grupo todas las niñas y niños que tienen entre cero y menos de seis años.
- En un segundo grupo los que tienen entre seis y menos de 12 años.
- En un tercer grupo las y los que tienen entre 12 y menos de 15 años.
- En un cuarto grupo las edades de las y los jóvenes que tienen entre 15 y menos de 18 años.
- En un quinto grupo, las y los jóvenes que tienen entre 18 y menos de 24 años de edad.
- Y un último grupo, la población mayor a 24 años”.

A partir de allí, puedes organizar la información que tienes sobre tu comunidad.

- Primero sería dividir en diferentes grupos como lo pide el presidente municipal.
- Para ello, integraría cada uno de los grupos de edad que mencionó el presidente municipal.
- Puedes proponer organizar los grupos por cada categoría que mencionó el presidente municipal por lo que serán 5 y uno más para incluir a todos los que no estaban en la división propuesta por el presidente municipal.
- Hay que considerar que los grupos deberían ser del mismo tamaño, así todos manejarían el mismo número de edades.

Población en edad preescolar y en edad típica de estudiar en mi comunidad (Distribución de frecuencias organizadas en clases)								
Clases			Cuadro de conteo	Frecuencia absoluta f_a	Frecuencia absoluta acumulada f_{aa}	Frecuencia relativa f_r	Frecuencia relativa acumulada f_{ra}	Marca de clase m_i
Límite inferior	-	Límite superior						
0	-	menos de 6	III	3	3	6	6	2.5
6	-	menos de 12	IIIIIIII	8	11	16	22	8.5
12	-	menos de 15	IIIIII	7	18	14	36	13
15	.	menos de 18	II	2	20	4	40	16
18	-	menos de 24	IIIIIIIIII	11	31	22	62	20.5
24	-	mayores de 24	IIIIIIIIII IIIIIIII	19	50	38	100	33.5
Suma total				50		100		

Actividad de aprendizaje 10

2. Sobre la observación de las gráficas se puede apreciar que:

- Gráfica 1. Esta gráfica nos presenta la información sobre el aumento de la obesidad en México para los años 1988, 1999 y 2006. No se especifica tipo de variable de medición. Fuente: Acuerdo Nacional para la Salud Alimentaria. 2010.
- Gráfica 2. Esta gráfica representa la población total por miles por grupo de edad y sexo en 2008 y 2012 de un estado de la República Mexicana. Los grupos de edad van desde 14 años hasta 75 o más años. Fuente: ECO-SALUD 2004.
- Gráfica 3. Esta gráfica representa la distribución del índice de masa corporal en alumnos de medicina de 1ro a 6to semestre. La fuente es la Secretaría de salud, 2012. y la variables de índice de masa corporal es categórica, se divide en desnutrición, normal, sobrepeso I, sobrepeso II y obesidad moderada.

4. La primera y la tercera son gráficas referentes a la salud. Todas hablan sobre poblaciones, aunque de diferente lugar y tamaño. La primera y la tercera no tienen diferencia de sexo.

Actividad de aprendizaje 11

Barra 1: 7.5

Barra 2: 68

Barra 3: 91

Barra 4: 32

Barra 5: 225.5

Barra 6: 636.5

Actividad de aprendizaje 12

Puedo decir que sí pues son comunidades que se parecen en el municipio, sin embargo, no es correcto porque faltan detalles estadísticos que no fueron considerados, la forma en la que se obtuvieron los resultados puede ser sesgada y poco confiable.

Para que una información obtenida de la muestra sea aplicable a todo el municipio debemos garantizar que nuestra muestra represente a toda la población es decir que en ella estén representadas todas las edades de las personas, que los grupos etarios de la muestra guarden la proporción con respecto a los formados en la población. Además la forma de seleccionar a las personas también debe de ser una en la que yo no las escoja sino sean elegidas para participar al azar.

Actividad de aprendizaje 1

- La estadística descriptiva es una rama de la estadística que nos permite obtener conclusiones sobre un conjunto de datos a través de la recolección, reorganización, presentación (tablas y gráficas) y también del análisis e interpretación de los datos. Esto nos permite observar más allá de los datos originales. Con la estadística descriptiva una lista de información, se convierte en frecuencias, tablas y gráficas, entre otros elementos que sirven para obtener información menos evidente.
- La diferencia se basa en el uso de estas ramas de la estadística, la estadística descriptiva se enfoca en la muestra y la estadística inferencial o inferencia estadística hace conclusiones a partir de las muestras referentes al comportamiento de la población.
- Si eligiera la primera opción, podría dar un número que me incluyera todas las edades y el número de personas que entrevisté, pero si tenemos muchas edades alejadas como pasaba en El Capulín, de tener edades cercanas a los 90 años, no considero que pudiera ser conveniente pues el número sólo se iría cercano a estas edades muy grandes y se prefería la variedad de la masa de datos.
- Si eligiera la segunda opción, podría dar un número en medio de un conjunto ordenado de datos, de un lado tendría la misma cantidad de datos que del otro lado de este número.
- Si eligiera la última opción, la edad que más se repite pues salta a la vista sin embargo deja de lado el resto de las edades que también dicen mucho sobre el conjunto que estudiamos.
- La edad más representativa puede ser la más grande porque es la máxima edad, puede ser la edad que represente mejor, sin embargo se pierde de nuevo la riqueza que hay en la variabilidad de los datos.

Actividad de aprendizaje 2

Seguramente habrás observado si los valores de los años están muy alejados o no del cálculo de la media y si existe un valor que hace este sesgo.

Actividad de aprendizaje 3

El Q1, corresponde al dato 25, el Q3 al dato 75, el D3 al 30, el D7 al 70 y el P40 al dato 40

Actividad de aprendizaje 4

El desempeño del módulo I puede interpretarse como bueno ya que tiene un valor superior al tercer cuartil, es decir superior al 75% del total. Sin embargo, excelente sería cercano al decil superior.

Actividades de aprendizaje 5 y 6

Población en edad preescolar y en edad típica de estudiar en mi comunidad (Distribución de frecuencias organizadas en clases)							
Clases	Intervalo	Número de observaciones	Marca de clase M_c	Frecuencia absoluta f_a	Frecuencia absoluta acumulada f_{aa}	Frecuencia relativa f_r	Frecuencia relativa acumulada f_{ra}
Los que no tienen edad de estudiar la primaria	0 a menos de 6	3	2.5	7.5	7.5	0.7	0.7
Los que tienen edad para estudiar la primaria	6 a menos de 12	8	8.5	68	75.5	6.6	7.3
Los que tienen edad de cursar la educación media (secundaria y bachillerato)	12 a menos de 18	9	14.5	130.5	206	12.7	20.0
Los que están en edad de cursar la universidad	18 a menos de 24	11	20.5	225.5	431.5	21.9	41.9
Los que no tienen edad para estudiar	24 a menos de 30	8	26.5	212	643.5	20.6	62.5
	30 a menos de 36	7	32.5	227.5	871	22.1	84.6
	36 o +	4	39.5	158	1029	15.4	100.0
Suma total		50		1029		100.0	

	$faacm$ = frecuencia acumulada hasta antes de la clase media
	fc_m = frecuencia de la clase media
	n = número de observaciones
	e = intervalo de clase
	Li = valor del límite inferior de la clase en la que cae la mediana

$$Me = 18 + \left[\frac{\left(\frac{50}{2} \right) - 20}{11} \right] 6$$

$$Me = 18 + \left[\frac{25 - 20}{11} \right] 6$$

$$Me = 18 + \left[\frac{5}{11} \right] 6$$

$$Me = 18 + [.45]6$$

$$Me = 18 + 2.7$$

$$Me = 20.7$$

Actividad de aprendizaje 8

Esta división en cuartiles de la población nos permite hacer conclusiones en diferentes estratos, es decir, nos permite ver el comportamiento de la población en fragmentos por tanto se espera que el alumno elabore una descripción estadística sobre las edades de cada uno de los cuartiles.

$$Q_1 = 12 + \left[\frac{(50/4) - 11}{9} \right] 6$$

$$Q_1 = 12 + \left[\frac{12.5 - 11}{9} \right] 6$$

$$Q_1 = 12 + \left[\frac{1.5}{9} \right] 6$$

$$Q_1 = 12 + [.16]6$$

$$Q_1 = 12 + .96$$

$$Q_1 = 12.96$$

$$Q_2 = 18 + \left[\frac{(50/4 \cdot 2) - 20}{11} \right] 6$$

$$Q_2 = 18 + \left[\frac{(12.5 * 2) - 20}{11} \right] 6$$

$$Q_2 = 18 + \left[\frac{25 - 20}{11} \right] 6$$

$$Q_2 = 18 + \left[\frac{5}{11} \right] 6$$

$$Q_2 = 18 + [.45]6$$

$$Q_2 = 18 + 2.7$$

$$Q_2 = 20.7$$

$$Q_3 = 24 + \left[\frac{(50/4 \cdot 3) - 31}{8} \right] 6$$

$$Q_3 = 24 + \left[\frac{(12.5 * 3) - 31}{8} \right] 6$$

$$Q_3 = 24 + \left[\frac{37.5 - 31}{8} \right] 6$$

$$Q_3 = 24 + \left[\frac{6.5}{8} \right] 6$$

$$Q_3 = 24 + [.81]6$$

$$Q_3 = 24 + 4.86$$

$$Q_3 = 28.86$$

Actividad de aprendizaje 9

$$D_1 = 6 + \left[\frac{(50/10 \cdot 1) - 3}{8} \right] 6$$

$$D_1 = 6 + \left[\frac{(5 \cdot 1) - 3}{8} \right] 6$$

$$D_1 = 6 + \left[\frac{2}{8} \right] 6$$

$$D_1 = 6 + [.25]6$$

$$D_1 = 6 + 1.5$$

$$D_1 = 7.5$$

$$D_6 = 18 + \left[\frac{(50/10 \cdot 6) - 20}{11} \right] 6$$

$$D_6 = 18 + \left[\frac{(5 \cdot 6) - 20}{11} \right] 6$$

$$D_6 = 18 + \left[\frac{30 - 20}{11} \right] 6$$

$$D_6 = 18 + \left[\frac{10}{11} \right] 6$$

$$D_6 = 18 + [.90]6$$

$$D_6 = 18 + 5.4$$

$$D_6 = 23.4$$

$$P_{23} = 12 + \left[\frac{(50/100 \cdot 23) - 11}{9} \right] 6$$

$$P_{23} = 12 + \left[\frac{(.5 \cdot 23) - 11}{9} \right] 6$$

$$P_{23} = 12 + \left[\frac{(11.5 - 11)}{9} \right] 6$$

$$P_{23} = 12 + \left[\frac{.5}{9} \right] 6$$

$$P_{23} = 12 + [0.05]6$$

$$P_{23} = 12 + 0.3$$

$$P_{23} = 12.3$$

$$P_{64} = 24 + \left[\frac{(50/100 \cdot 64) - 31}{8} \right] 6$$

$$P_{64} = 24 + \left[\frac{(.5 \cdot 64) - 31}{8} \right] 6$$

$$P_{64} = 24 + \left[\frac{(32 - 31)}{8} \right] 6$$

$$P_{64} = 24 + \left[\frac{1}{8} \right] 6$$

$$P_{64} = 24 + [.12]6$$

$$P_{64} = 24 + .72$$

$$P_{64} = 24.7$$

Actividad de aprendizaje 10

La moda como se ha explicado es el número que más se repite en un conjunto de datos, por lo tanto se espera ver que se enfatice en el número con más repeticiones. En el caso de datos agrupados que tienen la misma amplitud, se destacará el número y la clase modal que lo contiene como la de las mayores repeticiones.

$$M^o = 18 + \left(\frac{2}{2+3}\right)6$$

$$M^o = 18 + \left(\frac{2}{5}\right)6$$

$$M^o = 18 + (.4)6$$

$$M^o = 18 + 2.4$$

$$M^o = 20.4$$

Actividad de aprendizaje 11

\bar{x} = media aritmética = $147/10=14.7$

Diferencia entre cada dato - \bar{X}		
	$X_{ij} - \bar{X}$	$(X_{ij} - \bar{X})^2$
22-14.7=	7.3	53.29
9-14.7=	-5.7	32.49
8-14.7=	-6.7	44.89
13-14.7=	-1.7	2.89
15-14.7=	.3	0.09
24-14.7=	9.3	86.49
30-14.7=	15.3	134.09
14-14.7=	-.7	0.49
9-14.7=	-5.7	32.49
3-14.7=	-11.7	138.89
	64.4*	626.1

* La suma de las diferencias deberá darte cero. Si no tomas en cuenta el signo éste sería el total.

$$\text{Desviación media} = \frac{64.4}{10} = 6.44 = 6.44$$

$$\text{Desviación estándar} = \frac{\sqrt{626.1}}{10} = \sqrt{62.61} = 7.9$$

$$\text{Varianza} = 62.61$$

Actividad de aprendizaje 12

Población en edad preescolar y en edad típica de estudiar en mi comunidad (Distribución de frecuencias organizadas en clases)									
Clases	Intervalo	Número de observaciones	Marca de clase M_c	Frecuencia absoluta f_a	Frecuencia absoluta acumulada f_{aa}	Frecuencia relativa f_r	Frecuencia relativa acumulada f_{ra}	$2 - x$	$7 * 1$
Los que no tienen edad de estudiar la primaria	0 a menos de 6	3	2.5	7.5	7.5	0.7	0.7	-18.08	-54.24
Los que tienen edad para estudiar la primaria	6 a menos de 12	8	8.5	68	75.5	6.6	7.3	-12.08	-96.64
Los que tienen edad de cursar la educación media (secundaria y bachillerato)	12 a menos de 18	9	14.5	130.5	206	12.7	20.0	-6.08	-54.72
Los que están en edad de cursar la universidad	18 a menos de 24	11	20.5	225.5	431.5	21.9	41.9	-0.08	-0.88
	24 a menos de 30	8	26.5	212	643.5	20.6	62.5	5.92	47.36
Los que no tienen edad para estudiar	30 a menos de 36	7	32.5	227.5	871	22.1	84.6	11.92	83.44
	36 o +	4	39.5	158	1029	15.4	100.0	18.92	75.68
Suma total		50		1029		100.0			0

Desviación media =8.25

Actividad de aprendizaje 13

Población en edad preescolar y en edad típica de estudiar en mi comunidad (Distribución de frecuencias organizadas en clases)											
Clases	Intervalo	Número de observaciones	Marca de clase M_c	Frecuencia absoluta f_a	Frecuencia absoluta acumulada f_{aa}	Frecuencia relativa f_r	Frecuencia relativa acumulada f_{ra}	$2 - x$	$7 * 1$	7^2	$9 * 1$
Los que no tienen edad de estudiar la primaria	0 a menos de 6	3	2.5	7.5	7.5	0.7	0.7	-18.08	-54.24	326.9	980.7
Los que tienen edad para estudiar la primaria	6 a menos de 12	8	8.5	68	75.5	6.6	7.3	-12.08	-96.64	145.9	1167.4
Los que tienen edad de cursar la educación media (secundaria y bachillerato)	12 a menos de 18	9	14.5	130.5	206	12.7	20.0	-6.08	-54.72	37.0	332.7
Los que están en edad de cursar la universidad	18 a menos de 24	11	20.5	225.5	431.5	21.9	41.9	-0.08	-0.88	0.0	0.1
Los que no tienen edad para estudiar	24 a menos de 30	8	26.5	212	643.5	20.6	62.5	5.92	47.36	35.0	280.4
	30 a menos de 36	7	32.5	227.5	871	22.1	84.6	11.92	83.44	142.1	994.6
	36 o +	4	39.5	158	1029	15.4	100.0	18.92	75.68	358.0	1431.9
	Suma total	50		1029		100.0			0		5187.7

Suma columna 10 = 5187.7
Raíz cuadrada = 72.03
Divido entre el total de personas = 1.4404
Desviación estándar = 1.44
Varianza = 103.754

Actividad de aprendizaje 1

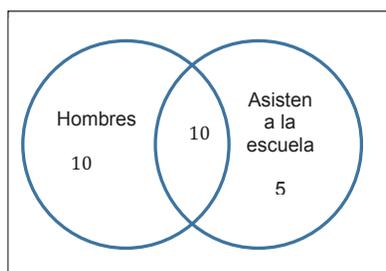
- Pato
- El loro y el pingüino son aves
- El loro y el pez volador son seres vivos que vuelan
- El pingüino y el pez volador son seres vivos que nadan

No tienen similitudes, porque el avestruz es un ave, la mariposa un ser vivo que vuela y la ballena un ser vivo que nada. El diagrama te ayuda de manera gráfica y simple a identificar los elementos con características similares y saber cuáles son las similitudes.

Actividad de aprendizaje 2

- Ballena es elemento de C
- Loro pertenece a A y B
- Avestruz \notin B y C
- Pez volador \notin A

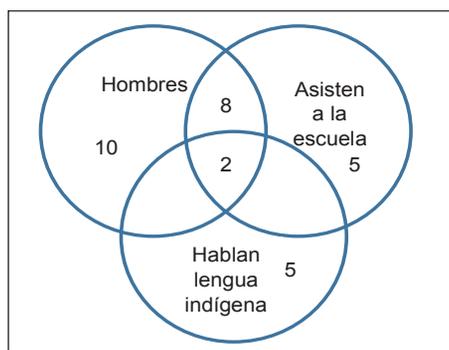
Actividad de aprendizaje 3



Para incorporar a los hombres que hablan lengua indígena agregaría un tercer conjunto que interseccione los dos anteriores y se logre la intersección de hombres + asisten a la escuela + hablan lengua indígena.

Actividad de aprendizaje 4

La característica que tienen en común los diagramas anteriores es que ambos representan la intersección de dos conjuntos, sin embargo en el segundo diagrama, un conjunto está contenido en el otro (el conjunto de las Madres está contenido dentro del conjunto de Mujeres):



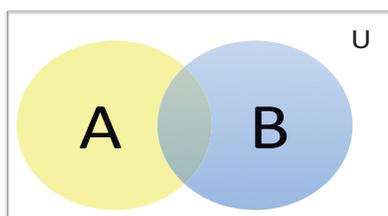
Actividad de aprendizaje 5

1. Niños menores de 12 a 18
2. Conjunto de personas mayores a
3. Ninguno

Actividad de aprendizaje 6

Unión	Intersección
Soleado o nublado	Hombres y solteros
Alto o bajo	Mayores de edad y abuelos
Día o noche	Peces y voladores

Actividad de aprendizaje 7



Actividad de aprendizaje 8

Primer diagrama. Representa los que no hablan alguna lengua indígena
 Segundo diagrama. Representa a los que no son hombres
 Tercer diagrama. Representa a los que no son hombres

Actividad de aprendizaje 11

68/87

El número de canicas que son verdes son 68 y el número de canicas totales en la bolsa son 87.

Actividad de aprendizaje 12

3. No, ya que hay más mujeres que hombres y por lo tanto no es igual la probabilidad de elegir una mujer que un hombre

Actividad de aprendizaje 13

1. La ventaja es que ya cuento con esa información trabajada estadísticamente, y la desventaja es que generalizar la información no necesariamente es conveniente pues el comportamiento de las edades de la población puede ser diferente al de la comunidad.

2. La ventaja es que tendríamos toda la información de toda la población, y la desventaja es que esto tomará mucho tiempo y es difícil encontrar a todos los habitantes de la comunidad.
3. La ventaja es que es información real, pero puede que no esté actualizada.

Actividad de aprendizaje 14

1. Hay hombres y mujeres
Hay varios grupos de edad
Algunos hablan español, otros hablan lengua indígena
Algunos asisten a la escuela
Algunos hacen deporte
2. Asegurando que la muestra sea aleatoria, pero que represente los distintos sectores y características de la población.
3. Elegiría habitantes a lo largo de todo el territorio para no centrarme solamente en una comunidad, de esta forma obtendría una muestra representativa.

Actividad de aprendizaje 15

50/2500

pues con este método de selección, todos los habitantes tienen la misma probabilidad de ser elegidos para entrevista.

Bibliográficas

Chao, L. (2002). Introducción a la estadística. México: McGraw-Hill.

Elorza, H. (2009). Estadística para las ciencias sociales, del comportamiento y de la salud. México: Cengage Learning.

Freud, J. y Simon, G. (1992). Estadística elemental. México: Prentice Hall.

Levin, R. I. y Rubin D. S. (1996). Estadística para Administradores. México: Prentice Hall.

Mendenhall, W. y Scheaffer, R. (2002). Estadística aplicada. México: Thomson International.

Méndez, I., Eslava, G. y Romero, P. (2004). Conceptos básicos de muestreo. Monografías vol. 12, No. 27., Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas. Universidad Nacional Autónoma de México.

Spiegel, M. (2003). Probabilidad y Estadística. México: McGraw-Hill.

Sitios web

<http://www.ub.edu/stat/GrupsInnovacio/Statmedia/demo/Temas/Capitulo6B0C6m1t9.htm>

<http://www.universoformulas.com/estadistica/inferencia/muestreo-estratificado/>

http://www.vitutor.com/pro/2/a_e.html

<http://frecuenciaestadistica.blogspot.com/2009/04/definicion-de-estadistica.html>

http://www.ugr.es/~rescate/practicum/el_m_todo_de_observaci_n.htm

http://ensanut.insp.mx/doctos/FactSheet_ResultadosNacionales14Nov.pdf









INSERTAR DATOS DE COLOFÓN

