



PROBABILIDAD Y ESTADÍSTICA I

GUÍA DIDÁCTICA DEL ESTUDIANTE





TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

Probabilidad y Estadística I

DATOS DEL ALUMNO

Nombre: _____ Plantel: _____

6to Semestre Grupo: _____ Turno: _____

Coloca en la tabla los días que tengas Probabilidad y estadística I y en qué número de clase del día

Clases en el día	LUNES	MARTES	MIÉRCOLES	JUEVES	VIERNES
1ª Clase					
2ª Clase					
3ª Clase					
4ª Clase					
5ª Clase					
6ª Clase					
7ª Clase					

Cobanota:

“

IMPORTANTE:

”

Directorio de COLEGIO DE BACHILLERES DE TABASCO

MTRO. ERASMO MARTÍNEZ RODRÍGUEZ

Director General

C.P. SONIA LÓPEZ IZQUIERDO

Directora Académico

DRA. GISELLE OLIVARES MORALES

Subdirectora de Planeación Académica

MTRA. ALEJANDRINA LASTRA COLORADO

Jefe de Departamento de Programas de Estudio

DR. REYLE MAR SARAO

Jefe de Materia

ASIGNATURA: PROBABILIDAD Y ESTADÍSTICA I.

Edición: 2021

En la realización del presente material, participaron:

Docente	Centro
Yahaira Méndez Sánchez	Plantel 1
José Eduardo Mendoza Pérez	Plantel 2
Luis Miguel Ruiz Rodríguez	Plantel 2
Diana Emily Peregrino Jiménez*	Plantel 3
José de Jesús Winzig	Plantel 4
Nancy Arias Chablé	Plantel 4
Juana Antonia Moo Salvador	Plantel 4
Lorenzo Mendoza Gómez*	Plantel 5
Ramón Augusto Escobar Priego	Plantel 5
Luis Felipe Córdova Carrasco	Plantel 6
Daniel Pérez Acosta	Plantel 7
Luis Francisco Gil Gaspar	Plantel 7
Seydi Guadalupe de la O Colomé*	Plantel 21
Enny Dolores Morales López	Plantel 22
Marcela Mendoza Sánchez	Plantel 28

***Asesor académico**



Este material fue elaborado bajo la coordinación y supervisión del Departamento de Programas de Estudio de la Dirección Académica del Colegio de Bachilleres de Tabasco, concluyendo en el mes de junio del 2021.

©Derechos en proceso de registro

Queda prohibida la reproducción total o parcial de este material por cualquier medio electrónico o mecánico, sin autorización del COBATAB.

Para uso interno de la Comunidad del Colegio de Bachilleres de Tabasco (COBATAB)

www.cobatab.edu.mx

DATOS DEL ALUMNO	2
PROBABILIDAD Y ESTADÍSTICA I	2
DIRECTORIO DE COLEGIO DE BACHILLERES DE TABASCO	3
PRESENTACIÓN	11
FUNDAMENTACIÓN	13
COMPETENCIAS GENÉRICAS	14
COMPETENCIAS DISCIPLINARES	18
ENFOQUE DE LA ASIGNATURA	19
UBICACIÓN DE LA ASIGNATURA	20
RELACIÓN DE LOS CONTENIDOS CON LOS APRENDIZAJES CLAVES	20
EVALUACIÓN POR COMPETENCIAS	21
INSTRUMENTOS DE EVALUACIÓN	22
TEMARIO	24
BLOQUE I: ELEMENTOS ESTADÍSTICOS	26
BLOQUE II: DESCRIPCIÓN GRÁFICA DE UN CONJUNTO DE DATOS	26
PROPÓSITO DE LOS BLOQUES I Y II	27
APRENDIZAJES ESPERADOS BLOQUE I Y II	27
SITUACIÓN DIDÁCTICA 1	29
PROPÓSITO DE LA SITUACIÓN DIDÁCTICA 1	29
INSTRUMENTO DE EVALUACIÓN SITUACIÓN DIDÁCTICA 1	30
EVALUACIÓN DIAGNÓSTICA SD1	31
¿QUÉ TANTO SÉ? (APERTURA)	31
DE LA INFORMACIÓN AL CONOCIMIENTO	32
¿QUÉ ES LA ESTADÍSTICA?	32
OBJETIVOS DE LA ESTADÍSTICA	32
DIVISIONES DE LA ESTADÍSTICA.	33
ELEMENTOS ESTADÍSTICOS	33
POBLACIÓN	33
MUESTRA	33
FRECUENCIA	34
DATO	34
TIPOS DE DATOS	34

VARIABLE _____	34
VARIABLE Y SUS TIPOS _____	34
TÉCNICAS DE RECOLECCIÓN DE DATOS _____	35
PRINCIPALES TÉCNICAS DE RECOLECCIÓN DE DATOS _____	35
ENCUESTA _____	36
ENTREVISTA _____	36
OBSERVACIÓN _____	36
EXPERIMENTACIÓN _____	36
DOCUMENTAL _____	36
TÉCNICAS DE MUESTREO _____	37
TIPOS DE MUESTREO _____	37
MUESTREO PROBABILÍSTICO _____	37
MUESTREO ALEATORIO SIMPLE _____	37
MUESTREO SISTEMÁTICO _____	38
MUESTREO ESTRATIFICADO _____	38
MUESTREO POR CONGLOMERADOS _____	39
MUESTREO NO PROBABILÍSTICO _____	40
MUESTREO POR CONVENIENCIA _____	41
MUESTREO BOLA DE NIEVE _____	41
MUESTREO POR CUOTAS _____	42
¿CUÁNDO USAR EL MUESTREO NO PROBABILÍSTICO? _____	43
TAREA NO. 1 _____	44
PYEI-B2-LC01 LISTA DE COTEJO PARA EVALUAR TAREA 01 CUADRO COMPARATIVO _____	46
DE LA INFORMACIÓN AL CONOCIMIENTO _____	47
DISTRIBUCIÓN DE FRECUENCIAS _____	47
ORGANIZACIÓN DE DATOS _____	47
DATOS NO AGRUPADOS _____	47
DATOS AGRUPADOS _____	47
TABLA DE DISTRIBUCIÓN DE FRECUENCIAS PARA DATOS AGRUPADOS _____	48
DETERMINACIÓN DE CLASES PARA DATOS NO AGRUPADOS _____	50
DISTRIBUCIÓN DE DATOS NO AGRUPADOS CON VARIABLE CUALITATIVA _____	50
DISTRIBUCIÓN DE DATOS NO AGRUPADOS CON VARIABLE CUANTITATIVA _____	51
DISTRIBUCIÓN DE DATOS AGRUPADOS CON VARIABLE CUANTITATIVA _____	53
DETERMINACIÓN DE CLASES PARA DATOS AGRUPADOS _____	55
LÍMITES DE LOS INTERVALOS DE CLASE PARA VARIABLES NUMÉRICAS DISCRETAS _____	57
CRITERIOS DE REPRESENTACIÓN Y ANÁLISIS DE GRÁFICAS _____	61
COMPARACIÓN DE DOS O MÁS GRUPOS _____	62

HISTOGRAMA	63
UN HISTOGRAMA ES UTILIZADO PARA:	64
POLÍGONO DE FRECUENCIA	66
CARACTERÍSTICAS DEL POLÍGONO DE FRECUENCIAS	67
GRÁFICA CIRCULAR	68
CLAVE DEL GRÁFICO CIRCULAR	68
VENTAJAS Y DESVENTAJAS DEL GRÁFICO CIRCULAR	68
<i>Ventajas</i>	68
<i>Desventajas</i>	69
EJEMPLOS DE GRÁFICOS CIRCULARES	69
¿COMO SE ELABORA UNA GRÁFICA CIRCULAR?	71
GRÁFICA OJIVA	75
TAREA NO. 2	77
PYEI-B2-LC02 LISTA DE COTEJO PARA EVALUAR TAREA 02 PROBLEMARIO	81
TAREA NO. 3	82
PYEI-B1_B2-PP01	83
CUESTIONARIO TIPO PLANEA BLOQUE I Y II	83
PYEI-B1_B2-MA01 MAPA DE APRENDIZAJE PARA EVALUAR LOS APRENDIZAJES ESPERADOS	86
REFERENCIAS BLOQUE I Y II	88
BLOQUE III:	91
MEDIDAS ESTADÍSTICAS	91
PROPÓSITO DEL BLOQUE III	92
APRENDIZAJES ESPERADOS BLOQUE III	92
SITUACIÓN DIDÁCTICA 2	94
PROPÓSITO DE LA SITUACIÓN DIDÁCTICA 2	94
INSTRUMENTO DE EVALUACIÓN SITUACIÓN DIDÁCTICA 2	95
ACTIVIDAD CONSTRUYE T	96
EVALUACIÓN DIAGNÓSTICA	98
¿QUÉ TANTO SÉ? (APERTURA)	98
DE LA INFORMACIÓN AL CONOCIMIENTO	101
CONCEPTOS BÁSICOS	101
MEDIDAS DE TENDENCIA CENTRAL	101
¿DATOS NO AGRUPADOS?	101
¿DATOS AGRUPADOS?	101
MEDIA	102

MEDIANA _____	102
MODA _____	103
SESGO _____	104
MEDIDAS DE TENDENCIA CENTRAL PARA DATOS NO AGRUPADOS _____	106
MEDIA _____	106
MEDIANA _____	106
MODA _____	107
MEDIDAS DE SESGO O ASIMETRÍA _____	107
MEDIA _____	108
MEDIANA _____	108
MODA _____	108
MEDIDAS DE SESGO O ASIMETRÍA _____	109
MEDIDAS DE TENDENCIA CENTRAL PARA DATOS AGRUPADOS _____	109
EJEMPLO 1 _____	109
Media _____	110
Mediana _____	110
<i>Fórmula para calcular la mediana de datos no agrupados:</i> _____	111
Moda _____	112
<i>Fórmula para calcular la moda de datos no agrupados:</i> _____	113
DE LA INFORMACIÓN AL CONOCIMIENTO _____	114
LÍMITES ESTADÍSTICOS Y MEDIDAS DE POSICIÓN _____	114
PERCENTILES: _____	114
CUARTILES: _____	114
DECILES: _____	115
DATOS NO AGRUPADOS. _____	115
<i>Ejemplo:</i> _____	115
PARA DATOS AGRUPADOS _____	116
<i>Percentiles:</i> _____	116
<i>Cuartiles:</i> _____	117
<i>Deciles:</i> _____	119
TAREA NO. 4 _____	121
PYEI-B2-LC003 LISTA DE COTEJO PARA EVALUAR TAREA 04 CRUCIGRAMA _____	124
DE LA INFORMACIÓN AL CONOCIMIENTO _____	125
MEDIDAS DE TENDENCIA CENTRAL Y DE DISPERSIÓN PARA DATOS SIMPLES _____	125
DATOS NO AGRUPADOS. _____	125
MEDIA: _____	126
MEDIANA: _____	126
MODA: _____	129
SESGO: _____	129
RANGO _____	130

DESVIACIÓN MEDIA _____	131
VARIANZA _____	132
DESVIACIÓN ESTÁNDAR _____	133
MEDIDAS DE TENDENCIA CENTRAL Y DE DISPERSIÓN PARA DATOS AGRUPADOS _____	134
DATOS AGRUPADOS. _____	134
MEDIA _____	134
MEDIANA _____	135
MODA _____	137
SESGO: _____	138
RANGO _____	139
DESVIACIÓN MEDIA _____	139
VARIANZA _____	140
DESVIACIÓN ESTÁNDAR _____	141
TAREA NO. 5 _____	142
PYEI-B2-LC004 LISTA DE COTEJO PARA EVALUAR TAREA 05 PROBLEMARIO _____	144
PYEI-B3-PP02 _____	145
CUESTIONARIO TIPO PLANEA BLOQUE III _____	145
PYEI-B3-MA02 MAPA DE APRENDIZAJE PARA EVALUAR LOS APRENDIZAJES ESPERADOS _____	148
REFERENCIAS BLOQUE III _____	149
BLOQUE IV: _____	150
COMPORTAMIENTO DE DOS VARIABLES _____	150
PROPÓSITO DEL BLOQUE IV _____	151
APRENDIZAJES ESPERADOS BLOQUE IV _____	151
SITUACIÓN DIDÁCTICA 3 _____	153
PROPÓSITO DE LA SITUACIÓN DIDÁCTICA 3 _____	153
INSTRUMENTO DE EVALUACIÓN SITUACIÓN DIDÁCTICA 3 _____	154
EVALUACIÓN DIAGNÓSTICA _____	156
¿QUÉ TANTO SÉ? (APERTURA) _____	156
DE LA INFORMACIÓN AL CONOCIMIENTO _____	158
CORRELACION LINEAL _____	158
DE LA INFORMACIÓN AL CONOCIMIENTO _____	162
REGRESION LINEAL _____	162
<i>Ejemplo:</i> _____	165
TAREA No. 6 _____	167
PYEI-B2-GO03 GUÍA DE OBSERVACIÓN PARA EVALUAR TAREA 06 EJERCICIOS _____	170
PYEI-B2-GO04 GUÍA DE OBSERVACIÓN PARA EVALUAR TAREA 07 PROBLEMARIO _____	172



TABASCO

"Educación que genera cambio"



COBATAB

COLEGIO DE BACHILLERES DE TABASCO

DE LA INFORMACIÓN AL CONOCIMIENTO	173
PRUEBA CHI CUADRADO.	173
PASOS PARA LA REALIZACIÓN DE UNA PRUEBA DE INDEPENDENCIA CHI CUADRADA	176
TABLA DE VALORES CRÍTICOS DE LA DISTRIBUCIÓN CHI CUADRADA	178
PYEI-B2-GO04 GUÍA DE OBSERVACIÓN PARA EVALUAR TAREA 08 PROBLEMARIO	181
PYEI4-B4-PP03	182
CUESTIONARIO TIPO PLANEA BLOQUE IV	182
PYEI-B4-MA01 MAPA DE APRENDIZAJE PARA EVALUAR LOS APRENDIZAJES ESPERADOS	185
REFERENCIAS BLOQUE IV	186
HIMNO COLEGIO	187
PORRA INSTITUCIONAL	188
COBACHITO	189

Presentación

En la búsqueda de estrategias para el fortalecimiento del desarrollo de competencias tanto en la enseñanza del docente como en aprendizaje de los estudiantes y con la finalidad de homogenizar el lenguaje académico en el desarrollo de las planeaciones didácticas de las diversas asignaturas que conforman el cuarto semestre de la EMS regidas por la DGB. La Dirección General del Colegio de Bachilleres de Tabasco, a través de la participación de docentes del área de matemáticas adscritos a diferentes planteles, se ha dado a la tarea de aprovechar la potencialidad en la experiencia de la enseñanza de las matemáticas tanto en el aula, como en actividades extramuros y ha desarrollado esta guía para el estudiante que facilite a la vez el trabajo docente de MATEMÁTICAS IV. En ella se señalan los aspectos curriculares propios de la asignatura mostrando la distribución de los diferentes bloques que la conforman, relacionados con los aprendizajes claves, así como las competencias genéricas y disciplinares básicas a desarrollar. De acuerdo con el propósito de cada uno de ellos y a los aprendizajes esperados se muestra por cada bloque(s) una con la situación didáctica (SD) como problemática a resolver una vez abordados los contenidos específicos establecidos en los contenidos conceptuales.

Para el desarrollo de esta asignatura se han establecido **3 Situaciones Didácticas (SD)** seguidas de su instrumento de evaluación, con indicadores alineados a sus contenidos conceptuales, procedimentales y actitudinales que abonan al logro de los aprendizajes esperados al concluir el o los bloques que engloba dicha situación didáctica.

En la enseñanza bajo el enfoque por competencias se busca que los estudiantes adquieran aprendizajes que sean profundos, situados, significativos y socioemocionales, mismos que deben reflejarse en la solución de la problemática establecida en el conflicto cognitivo de la SD, por ello esta guía se proponen también **8 tareas** como parte de los insumos para la elaboración del producto que materialice los resultados del logro de los aprendizajes esperados en las **3 SD**, que se plantean en la planeación didáctica estatal. Dichas tareas también están acompañadas con su respectivo instrumento de evaluación.

Es importante mencionar que las tareas establecidas para cada bloque deben ser agotadas para dar paso a la presentación, socialización y evaluación del producto que a través de la estrategia nombrada da solución a cada situación didáctica.

En la planeación didáctica estatal se proponen los tipos de evaluaciones en las diversas tareas y situaciones didácticas, pero el docente tiene la libertad de elegir entre **autoevaluar**, **coevaluar** y **heteroevaluar** de acuerdo con los momentos del proceso de enseñanza-aprendizaje y del contexto de su grupo(s), lo importante es ejercer la práctica de evaluar; pues fortalece el proceso socio formativo en el aprendizaje de los estudiantes. Al final de cada sección que abarca cada bloque y sus respectivas situaciones didácticas y tareas se incluye un mapa de aprendizaje, esto para realizar una autoevaluación que permite a cada estudiante y al docente mismo conocer el nivel de logro en los aprendizajes establecidos para así diseñar un plan de mejora de las actividades de enseñanza-aprendizaje.

Para fortalecer el desarrollo del aprendizaje socioemocional se integra la lección CONSTRUYE-T a desarrollar, diferente a la que se aplica en las otras asignaturas del cuarto semestre.

Por último, no puede omitirse señalar que para facilitar el desarrollo de estrategias de trabajo en algunos contenidos en el aula y fuera de ella, se insertan códigos QR e imágenes con sus respectivo enlace o dirección electrónica, dándole la versatilidad que se requiere para la aplicación de aula invertida y el aprendizaje autónomo.

Este trabajo está alineado a la Planeación Didáctica de Matemáticas IV, en la que de acuerdo con las habilidades de cada docente puede aplicarse en modos de educación virtual y/o presencial. Esperamos fortalezca y facilite el desarrollo de su tarea educativa en esta asignatura

ATENTAMENTE

Docentes Participantes

Fundamentación

La Dirección General de Bachillerato ha iniciado la actualización de sus programas de estudio integrando elementos como: aprendizajes claves, contenidos específicos y aprendizajes esperados que atienden al Nuevo Modelo Educativo para la Educación Obligatoria. Además de conservar su enfoque en competencias hace énfasis en el desarrollo de las habilidades socioemocionales y aborda temas transversales tomando en cuenta lo estipulado en las políticas educativas vigentes.

Dicha actualización tiene fundamento en el Programa Sectorial de Educación 2013 -2018, en el que se señala que la Educación Media Superior debe ser fortalecida para contribuir en el desarrollo de México, formando en competencias a hombres y mujeres esenciales para construir una Nación próspera basada en el conocimiento.

Dicha actualización considera también la finalidad esencial del Bachillerato, que es generar en el estudiante el desarrollo de una primera síntesis personal y social que le permita su acceso a la educación superior, a la vez que le dé una comprensión de su sociedad y de su tiempo y lo prepare para una posible incursión en el sector productivo, así como también ofrecerle una cultura general básica que comprenda aspectos de las ciencias de las humanidades y de las técnicas para la construcción de nuevos conocimientos.

Siendo la asignatura parte del **Componente de Formación Propedéutico del Bachillerato General**, busca preparar al estudiantado para la continuación de sus estudios en escuelas diversas de educación superior.

Competencias genéricas

SE AUTODETERMINA Y CUIDA DE SÍ	
1. Se conoce y valora a sí mismo y aborda problemas y retos teniendo en cuenta los objetivos que persigue	
CG. 1.1.	Enfrenta las dificultades que se le presentan y es consciente de sus valores, fortalezas y debilidades.
CG. 1.2.	Identifica sus emociones, las maneja de manera constructiva y reconoce la necesidad de solicitar apoyo ante una situación que lo rebase.
CG. 1.3.	Elige alternativas y cursos de acción con base en criterios sustentados y en el marco de un proyecto de vida.
CG. 1.4.	Analiza críticamente los factores que influyen en su toma de decisiones.
CG. 1.5.	Asume las consecuencias de sus comportamientos y decisiones.
CG. 1.6.	Administra los recursos disponibles teniendo en cuenta las restricciones para el logro de sus metas.
2. Es sensible al arte y participa en la apreciación e interpretación de sus expresiones en distintos géneros	
CG 2.1	Valora el arte como manifestación de la belleza y expresión de ideas, sensaciones y emociones.
CG 2.2	Experimenta el arte como un hecho histórico compartido que permite la comunicación entre individuos y culturas en el tiempo y el espacio, a la vez que desarrolla un sentido de identidad.
CG 2.3	Participa en prácticas relacionadas con el arte.
3. Elige y practica estilos de vida saludables	
CG 3.1	Reconoce la actividad física como un medio para su desarrollo físico, mental y social.
CG 3.2	Toma decisiones a partir de la valoración de las consecuencias de distintos hábitos de consumo y conductas de riesgo.
CG 3.3	Cultiva relaciones interpersonales que contribuyen a su desarrollo humano y el de quienes lo rodean.



SE EXPRESA Y COMUNICA

4. Escucha, interpreta y emite mensajes pertinentes en distintos contextos mediante la utilización de medios, códigos y herramientas apropiados.

CG 4.1	Expresa ideas y conceptos mediante representaciones lingüísticas, matemáticas o gráficas.
CG 4.2	Aplica distintas estrategias comunicativas según quienes sean sus interlocutores, el contexto en el que se encuentra y los objetivos que persigue.
CG 4.3	Identifica las ideas clave en un texto o discurso oral e infiere conclusiones a partir de ellas.
CG 4.4	Se comunica en una segunda lengua en situaciones cotidianas.
CG 4.5	Maneja las tecnologías de la información y la comunicación para obtener información y expresar ideas.

5. Desarrolla innovaciones y propone soluciones a problemas a partir de métodos establecidos

CG 5.1	Sigue instrucciones y procedimientos de manera reflexiva, comprendiendo como cada uno de sus pasos contribuye al alcance de un objetivo.
CG 5.2	Ordena información de acuerdo con categorías, jerarquías y relaciones.
CG 5.3	Identifica los sistemas y reglas o principios medulares que subyacen a una serie de fenómenos.
CG 5.4	Construye hipótesis y diseña y aplica modelos para probar su validez.
CG 5.5	Sintetiza evidencias obtenidas mediante la experimentación para producir conclusiones y formular nuevas preguntas.
CG 5.6	Utiliza las tecnologías de la información y comunicación para procesar e interpretar información.

6. Sustenta una postura personal sobre temas de interés y relevancia general, considerando otros puntos de vista de manera crítica y reflexiva

CG 6.1	Elige las fuentes de información más relevantes para un propósito específico y discrimina entre ellas de acuerdo a su relevancia y confiabilidad.
CG 6.2	Evalúa argumentos y opiniones e identifica prejuicios y falacias.
CG 6.3	Reconoce los propios prejuicios, modifica sus puntos de vista al conocer nuevas evidencias, e integra nuevos conocimientos y perspectivas al acervo con el que cuenta.
CG 6.4	Estructura ideas y argumentos de manera clara, coherente y sintética.

APRENDE DE FORMA AUTÓNOMA



7. Aprende por iniciativa e interés propio a lo largo de la vida

CG 7.1	Define metas y da seguimiento a sus procesos de construcción de conocimiento.
CG 7.2	Identifica las actividades que le resultan de menor y mayor interés y dificultad, reconociendo y controlando sus reacciones frente a retos y obstáculos.
CG 7.3	Articula saberes de diversos campos y establece relaciones entre ellos y su vida cotidiana.

TRABAJA EN FORMA COLABORATIVA

8. Participa y colabora de manera efectiva en equipos diversos

CG 8.1	Propone maneras de solucionar un problema o desarrollar un proyecto en equipo, definiendo un curso de acción con pasos específicos.
CG 8.2	Aporta puntos de vista con apertura y considera los de otras personas de manera reflexiva.
CG 8.3	Asume una actitud constructiva, congruente con los conocimientos y habilidades con los que cuenta dentro de distintos equipos de trabajo.

PARTICIPA CON RESPONSABILIDAD EN LA SOCIEDAD

9. Participa con una conciencia cívica y ética en la vida de su comunidad, región, México y el mundo

CG 9.1	Privilegia el diálogo como mecanismo para la solución de conflictos.
CG 9.2	Toma decisiones a fin de contribuir a la equidad, bienestar y desarrollo democrático de la sociedad.
CG 9.3	Conoce sus derechos y obligaciones como mexicano y miembro de distintas comunidades e instituciones, y reconoce el valor de la participación como herramienta para ejercerlos.
CG 9.4	Contribuye a alcanzar un equilibrio entre el interés y bienestar individual y el interés general de la sociedad.
CG 9.5	Actúa de manera propositiva frente a fenómenos de la sociedad y se mantiene informado.
CG 9.6	Advierte que los fenómenos que se desarrollan en los ámbitos local, nacional e internacional ocurren dentro de un contexto global interdependiente.

10. Mantiene una actitud respetuosa hacia la interculturalidad y la diversidad de creencias, valores, ideas y prácticas sociales

CG 10.1	Reconoce que la diversidad tiene lugar en un espacio democrático de igualdad de dignidad y derechos de todas las personas, y rechaza toda forma de discriminación.
CG 10.2	Dialoga y aprende de personas con distintos puntos de vista y tradiciones culturales mediante la ubicación de sus propias circunstancias en un contexto más amplio.



CG 10.3	Asume que el respeto de las diferencias es el principio de integración y convivencia en los contextos local, nacional e internacional.
11. Contribuye al desarrollo sustentable de manera crítica, con acciones responsables	
CG 11.1	Asume una actitud que favorece la solución de problemas ambientales en los ámbitos local, nacional e internacional.
CG 11.2	Reconoce y comprende las implicaciones biológicas, económicas, políticas y sociales del daño ambiental en un contexto global interdependiente.
CG 11.3	Contribuye al alcance de un equilibrio entre los intereses de corto y largo plazo con relación al ambiente.

Competencias disciplinares

CLAVE	MATEMÁTICAS
CDBM 1	Construye e interpreta modelos matemáticos mediante la aplicación de procedimientos aritméticos, algebraicos, geométricos y variaciones para la comprensión y análisis de situaciones reales, hipotéticas o formales.
CDBM 2	Formula y resuelve problemas matemáticos aplicando diferentes enfoques.
CDBM 3	Explica e interpreta los resultados obtenidos mediante procedimientos matemáticos y los contrasta con modelos establecidos o situaciones reales.
CDBM 4	Argumenta la solución obtenida de un problema, con métodos numéricos, gráficos, analíticos o variacionales, mediante el lenguaje verbal, matemático y el uso de las tecnologías de la información y la comunicación.
CDBM 5	Analiza las relaciones entre dos o más variables de un proceso social o natural para determinar o estimar su comportamiento.
CDBM 6	Cuantifica, representa y contrasta experimental o matemáticamente las magnitudes del espacio y las propiedades físicas de los objetos que lo rodean.
CDBM 7	Elige un enfoque determinista o uno aleatorio para el estudio de un proceso o fenómeno y argumenta su pertinencia.
CDBM 8	Interpreta tablas, graficas, mapas, diagramas y textos con matemáticos y científicos. símbolos

CONÓCENOS
COBATAB



Enfoque de la asignatura

El campo de Matemáticas tiene como eje desarrollar el pensamiento lógico- matemático, para interpretar situaciones reales o hipotéticas que permitan al estudiantado proponer alternativas de solución desde diversos enfoques ,priorizando las habilidades del pensamiento tales como la búsqueda de patrones o principios que subyacen a fenómenos cotidianos ,la generación de diversas alternativas para la solución de problemas, el manejo de información ,la toma de decisiones basados en el análisis crítico, la interpretación de tablas ,graficas ,textos, símbolos matemáticos que se encuentren en su entorno, y le permitirán la argumentación de una propuesta de solución o la predicción del comportamiento de un fenómeno a partir del análisis de sus variables

La asignatura de Probabilidad y estadística I con ENFOQUE EN ESTADISTICA tiene como propósito general ,desarrollar el pensamiento lógico matemático mediante el manejo de la información de situaciones de su contexto que le permitan aplicar la estadística descriptiva e inferencial como herramienta para la solución ,proyección y pronóstico de eventos en los cuales la toma de decisiones tiene injerencia en su vida.

Ubicación de la asignatura

1er Semestre	2do Semestre	3er Semestre	4to Semestre	5to Semestre	6to Semestre
Matemáticas I	Matemáticas II	Matemáticas III	Matemáticas IV	Geografía	Ecología y Medio Ambiente
Química I	Química II	Bilología I	Biología II	Estructura Socioeconómica de México	Historia Universal y Contemporánea
Taller de Lectura y Redacción I	Taller de Lectura y Redacción II	Física I	Física II	Probabilidad y Estadística I	Probabilidad y Estadística II
Ética y Valores I	Ética y Valores II	Todas las asignaturas de 3er semestre	Todas las asignaturas de 4° semestre	Se retomarán las asignaturas que en cada Plantel se imparten en 5to semestre, tanto del componente de formación propedéutico como de formación para el trabajo	Se retomarán las asignaturas que en cada Plantel se imparten en 6to semestre, tanto del componente de formación propedéutico como de formación para el trabajo
Metodología de la Investigación					
Informática I	Informática II				
Todas las asignaturas de 1er semestre	Todas las asignaturas de 2° semestre	FORMACIÓN PARA EL TRABAJO			
TUTORÍAS					

Relación de los contenidos con los aprendizajes claves

EJE	COMPONENTE	CONTENIDO CENTRAL	BLOQUE
Del manejo de la información al pensamiento estadístico	Riesgo, inferencia y aleatoriedad: Elementos de la estadística y la probabilidad	-Conceptos básicos de estadística y probabilidad	I
		-Concepto de riesgo en situaciones contextuales	II
		-Recolección de datos y su clasificación en clases	III
		-Manejo de la información en situaciones de la vida cotidiana.	IV
		-Tratamiento y significado de las medidas de tendencia central.	
		-Tratamiento y significado de las medidas de dispersión	
		-Uso del conteo y la probabilidad para eventos	
		-Contextualización de los elementos de probabilidad condicional e interpretación del Teorema de Bayes	

Evaluación por competencias

La evaluación debe ser un proceso continuo que permita recabar evidencias pertinentes sobre el logro de aprendizaje del estudiantado tomando en cuenta la diversidad de estilos y ritmos, con el fin de retroalimentar el proceso de enseñanza-aprendizaje y mejorar sus resultados.

El Modelo Educativo para la Educación Obligatoria ((MEPEO, sept 2017) señala que la evaluación es un proceso que tiene como objetivo mejorar el desempeño del alumnado e identificar sus áreas de oportunidad. Además, es un factor que impulsa la transformación de la práctica pedagógica y el seguimiento de los aprendizajes.

Para que la evaluación sea un proceso transparente y participativo donde se involucre al personal docente y al estudiantado debe favorecerse:

- **La autoevaluación**

En esta el bachiller valora sus capacidades con base a criterios y aspectos definidos con claridad por el personal docente el cual debe motivarle a buscar que tome conciencia de sus logros, errores y aspectos a mejorar durante su aprendizaje.

- **La coevaluación**

A través de la cual las personas pertenecientes al grupo valoran, evalúan y realimentan a un integrante en particular respecto a la presentación de evidencias de aprendizaje con base en criterios consensuados e indicadores previamente establecidos.

- **La heteroevaluación**

La cual consiste en un juicio emitido por el personal docente sobre las características del aprendizaje del estudiantado señalando las fortalezas y aspectos a mejorar teniendo como evidencia los aprendizajes logrados y evidencias específicas

Para evaluar por competencias se debe favorecer el proceso de formación a través de:

- **La evaluación diagnóstica**

Se realiza antes de algún proceso educativo (curso, secuencia, bloque de asignatura) para estimar los conocimientos previos del estudiantado, identificar sus capacidades cognitivas con relación al objeto de estudio y apoya al personal docente en la toma de decisiones del trabajo en el aula.

▪ La evaluación formativa

Se lleva a cabo durante el proceso educativo y permite precisar los avances logrados en el desarrollo de competencias por cada estudiante y advierte las dificultades que encuentra durante el aprendizaje. Tiene por objeto mejorar, corregir o reajustar su avance y se fundamenta en parte en la autoevaluación. Implica una reflexión y un diálogo con el estudiantado acerca de los resultados obtenidos y los procesos de aprendizaje y enseñanza que le llevaron a ello, permite estimar la eficacia de las experiencias de aprendizajes para mejorarlas y favorece su autonomía.

▪ La evaluación sumativa

Se realiza al final de un proceso o ciclo educativo considerando un conjunto de diversas evidencias que surgen de los aprendizajes logrados. Su fin consiste en certificar el grado en que las intenciones educativas se han alcanzado.

Instrumentos de evaluación

Con el fin de mostrar el saber que subyace en una competencia, los aprendizajes esperados permiten establecer una estrategia de evaluación, por tanto, contienen elementos observables que deben ser considerados en la evaluación tales como:

- La participación
- Las actividades generativas
- Las actividades de análisis

Para ello se consideran instrumentos que pueden agruparse principalmente en (Díaz-Barriga, 2014)

Técnicas de observación

- **Guía de observación:** Las técnicas de observación permiten evaluar los procesos de aprendizaje en el momento que se producen, La guía de observación es un instrumento que se basa en una lista de indicadores que pueden redactarse ya sea como afirmaciones o bien como preguntas, que orientan el trabajo de observación dentro del aula, señalando los aspectos que son relevantes al observar. Esta guía puede utilizarse para observar las respuestas de los alumnos en una actividad, durante una semana de trabajo, una secuencia didáctica completa.



"Educación que genera cambio"

Técnicas para el análisis del desempeño

- **Rúbricas**: Son guías que describen las características específicas de lo que se pretende evaluar (productos, tareas, proyectos, exposiciones, entre otras) precisando los niveles de rendimiento que permiten evidenciar los aprendizajes logrados de cada estudiante, valorar su ejecución y facilitar la retroalimentación.
- **Portafolios**: permiten mostrar el crecimiento gradual y los aprendizajes logrados con relación al programa de estudios, centrándose en la calidad o nivel de competencia alcanzado y no en una mera colección al azar de trabajos sin relación. Estos establecen criterios y estándares para elaborar diversos instrumentos para la evaluación del aprendizaje ponderando aspectos cualitativos de lo cuantitativo.
- **Listas de cotejo**: Es una lista de palabras, frases u oraciones que señalan con precisión las tareas, las acciones, los procesos y las actitudes que se desean evaluar

Los trabajos que pueden integrar en un portafolio y que pueden ser evaluados a través de rúbricas son: ensayos, videos, series de problemas resueltos, trabajos artísticos, trabajos colectivos, comentarios a lecturas realizadas, autorreflexiones, reportes de laboratorio, hojas de trabajo, guiones, entre otros, los cuales deben responder a una lógica de planeación o proyecto.

Con base a lo anterior, los programas de estudio de Dirección General del Bachillerato deben incluir elementos que enriquecen la labor formativa tales como la transversalidad, las habilidades socioemocionales y la interdisciplinariedad trabajadas de manera colegiada y permanentemente en el aula, consideran a la evaluación formativa como eje central al promover una reflexión sobre el progreso del desarrollo de competencias

Temario

BLOQUE	NOMBRE DEL BLOQUE	CONTENIDOS ESPECÍFICOS	HSM
I	Elementos estadísticos	<p>Elementos estadísticos</p> <ul style="list-style-type: none"> • Población • Dato • Muestra • Variable y sus tipos <p>Técnicas de recolección de datos</p> <ul style="list-style-type: none"> • Encuesta • Entrevista • Observación • Experimentación • Documental <p>Técnicas de muestreo</p> <ul style="list-style-type: none"> • Probabilístico • No probabilístico 	6
II	Descripción grafica de un conjunto de datos	<p>-Representación tabular mediante distribución de frecuencias</p> <ul style="list-style-type: none"> • Frecuencia absoluta • Frecuencia acumulada • Frecuencia relativa <p>-Determinación de clases para datos agrupados</p> <p>-Criterios, representación y análisis de graficas</p> <ul style="list-style-type: none"> • Histograma • Polígono de frecuencias • Circular • Ojiva 	12



BLOQUE	NOMBRE DEL BLOQUE	CONTENIDOS ESPECÍFICOS	HSM
III	Medidas estadísticas	<p>Medidas de tendencial central para datos simples y agrupados</p> <ul style="list-style-type: none">• Media• Mediana• Moda• Sesgo <p>Medidas de dispersión para datos simples y agrupados</p> <ul style="list-style-type: none">• Rango• Desviación media• Varianza• Desviación estándar <p>Limites estadísticos Cuartiles y percentiles</p>	10
IV	Comportamiento de dos variables	<p>Correlación de variables</p> <ol style="list-style-type: none">1. Representación2. Regresión lineal3. Ji-cuadrada	12



BLOQUE I: Elementos Estadísticos

BLOQUE II: Descripción Gráfica De Un Conjunto De Datos

Propósito de los bloques I y II



Bloque I

Emplea la estadística como herramienta para organizar, resumir y transmitir información, así como estimar comportamientos de su entorno que le permitan una mejor toma de decisiones, favoreciendo su pensamiento crítico.

Bloque II

Usa las representaciones tabulares y gráficas para comprender situaciones de su contexto, permitiéndole una toma de decisiones consciente e informada.

Aprendizajes Esperados

- Aplica crítica y reflexivamente los elementos estadísticos, así como la información recolectada a través de las diferentes técnicas, su tipo de variabilidad y su comportamiento en diferentes contextos



- Emplea las tablas de distribuciones de frecuencia para describir de manera crítica y reflexiva los resultados de investigaciones contextualizadas.
- Utiliza las gráficas como un medio creativo para comparar valores y facilitar la toma responsable de decisiones en problemas presentes en cualquier contexto

Competencias

Genéricas

3. Elige y practica estilos de vida saludable

3.2 Toma decisiones a partir de la valoración de las consecuencias de distintos hábitos de consumo y conductas de riesgo.

4. Escucha, interpreta y emite mensajes pertinentes en distintos contextos mediante la utilización de medios, códigos y herramientas apropiados

4.1 expresa ideas y conceptos mediante representaciones lingüísticas, matemáticas y gráficas

5. Desarrolla innovaciones y propone soluciones a problemas a partir de métodos establecidos

5.2 Ordena información de acuerdo con jerarquías, categorías y relaciones

8. participa y colabora de manera efectiva en equipos diversos

8.1 propone maneras de solucionar un problema o desarrollar un proyecto en equipo, definiendo un curso de acción con pasos específicos

8.2 aporta puntos de vista con apertura y considera los de otra persona de manera reflexiva

Disciplinares

2. Formula y resuelve problemas matemáticos aplicando diferentes enfoques

4. Argumenta la solución obtenida de un problema, con métodos gráficos analíticos o variacionales, mediante el lenguaje verbal, matemático y el uso de las tecnologías de la información y la comunicación

5. analiza las relaciones entre dos o más variables de un proceso social o natural para determinar o estimar su comportamiento.

7. elige un enfoque determinista o uno aleatorio para el estudio de un proceso o fenómeno y argumenta su pertinencia.

8. Interpreta tablas, gráficas, mapas, diagramas y textos con símbolos matemáticos y científicos.

Situación Didáctica 1

Título:

"Uniforma e informa"

Contexto:

La Dirección general de Colegio de Bachilleres de Tabasco, con el objetivo de apoyar la economía de las familias de sus estudiantes, pretende realizar un convenio de colaboración con una empresa de fabricantes de uniformes escolares (pantalones, playeras y faldas), por lo que cada plantel realizará un muestreo entre los alumnos de nuevo ingreso para conocer las tallas más comunes. Previendo que los fabricantes aprovechen al máximo su materia prima y abastecer a tiempo la demanda estudiantil del COBATAB.

Conflicto cognitivo:

- a) Describe la metodología que se debe seguir para conseguir un buen muestreo.
- b) ¿Cuál sería la mejor distribución de tallas en tu plantel?

Propósito de la situación didáctica 1

En equipos de 6 elementos **elaborar** las tablas y gráficas que representen las tallas de pantalones, faldas y playeras de los estudiantes de nuevo ingreso (muestra del 20% de la población), **mediante la recolección de datos y el muestreo para generar la** solución de la situación y presentarla ante el grupo para su socialización



Instrumento de Evaluación Situación Didáctica 1

PYE1-B1-B2-GO01 Guía de observación para evaluar la Situación Didáctica 1

Asignatura:	Probabilidad estadística I	y	<ul style="list-style-type: none"> • Bloque I • Bloque II 	<ul style="list-style-type: none"> • Elementos estadísticos • Descripción gráfica de un conjunto de datos 	Fecha:	
--------------------	----------------------------	---	---	---	---------------	--

Nombres	1.
	2.
	3.
	4.
	5.
	6.

Situación Didáctica 1:

Aprendizajes Esperados	Contenidos Específico
<ul style="list-style-type: none"> ▪ Aplica crítica y reflexivamente los elementos estadísticos, así como la información recolectada a través de las diferentes técnicas, su tipo de variabilidad y su comportamiento en diferentes contextos. ▪ Emplea las tablas de distribuciones de frecuencia para describir de manera crítica y reflexiva los resultados de investigaciones contextualizadas. ▪ Utiliza las gráficas como un medio creativo para comparar valores y facilitar la toma responsable de decisiones en problemas presentes en cualquier contexto. 	<ul style="list-style-type: none"> ▪ Elementos estadísticos ▪ Técnicas de recolección de datos ▪ Técnicas de muestreo ▪ Representación tabular mediante distribución de frecuencias ▪ Determinación de clases para datos agrupados ▪ Criterios, representación y análisis de graficas

CRITERIOS	%	CUMPLE		Puntaje
		SI	NO	
1. Entrega su producto terminado en el tiempo establecido por el facilitador	10			
2. Muestra el tipo de técnica utilizada para la investigación de la información	10			
3. Describe las ventajas que tiene el emplear diversas técnicas de muestreo para una misma población.	10			
4. Identifica las características de la población objetivo mediante la distribución de frecuencias reconociendo los pasos a seguir en la elaboración de tablas de distribución de frecuencias.	20			
5. Utiliza las tecnologías de la información para representar gráficamente las variables obtenidas en un histograma, polígono de frecuencias, circular u ojiva.	20			
6. Analiza la información obtenida para describir de manera crítica y reflexiva los resultados de la información obtenida para la adquisición de uniformes de DGB.	20			
7. Colabora con sus compañeros mostrando disposición al trabajo metódico y organizado.	10			
Calificación				

Logros obtenidos:

Aspectos por mejorar:

Evaluación Diagnóstica SD1

¿Qué tanto sé? (Apertura)



PYEI-B1-B2-ED01

DISCUSIÓN GUIADA

Bloque I "Elementos estadísticos". Bloque II "Descripción grafica de un conjunto de datos".

1. Partiendo de situaciones comunes de la vida cotidiana, enlista algunos casos en los que recuerdes que se emplea en términos como: estadística, población, muestra, muestreo y censo.

Estadística:

Población:

Muestra:

Muestreo:

Censo:

2. Es relativamente común ver en sitios públicos como calles, plazas o comercios a personas que se acercan a transeúntes para pedirles que respondan "una breve encuesta"
 - a) En términos generales, ¿en qué consiste una encuesta?
 - b) ¿Cuál crees que sea su propósito?
3. Es habitual que en los periódicos los medios de comunicación masiva publiquen una gran cantidad de datos sobre preferencias electorales obtenidos por medio de encuestas.
 - a) ¿De qué manera suelen presentar los resultados?
 - b) ¿Qué tipo de información presentan?
 - c) En tu opinión, ¿cuál es su utilidad?
4. Para ti, ¿qué es una variable?
5. Según lo que sabes, ¿qué es una tabla de datos estadísticos y para qué sirve?
6. ¿Para qué consideras que sirve un gráfico estadístico?
7. Enlista al menos 3 tipos de gráficos relacionados con la presentación de información estadística que alguna vez hayas visto y describe brevemente el contexto en el que aparecerían.
8. ¿Consideras que la presentación de información estadística es siempre objetiva?
¿Por qué?



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

De la información al conocimiento



Por mucho tiempo hemos tratado de predecir el futuro, como predecir si todo tu grupo pasará el semestre o no, hablando de un suceso próximo y cercano, pero hay cuestiones más importantes como determinar el porqué del calentamiento global y sus consecuencias, el crecimiento de la población y sus requerimientos alimenticios, vestido, espacio y servicios, lo que muestra que el hombre trata de saber lo que puede ocurrir en un futuro cercano y a largo plazo, ya que si no se evaluaran las posibilidades con anticipación, se podría crear una catástrofe mundial. Por ejemplo, las grandes compañías no pueden tomar decisiones basadas en un adivino que lee cartas o consulta su bola de cristal, para ellos es necesario tomar decisiones con un alto grado de seguridad y bases científicas certeras.

¿Qué es la estadística?

YOUTUBE:
Introducción a la
Estadística



<https://www.youtube.com/watch?v=UKzgs0fGXOE>

La estadística es la ciencia que se encarga de recolectar, describir e interpretar una cantidad de datos cuantitativos y cualitativos, los que se organizan (tabulan) y procesan para brindar información, tomar decisiones o inferir.

La estadística trabaja sobre una gran cantidad de datos, utiliza las bases de la matemática pura y las enlaza con el mundo real.

Cuando existe un fenómeno que puede ser analizado estadísticamente, podemos representarlo con un modelo matemático y por medio de este hacer predicciones futuras, mientras más datos se tengan, más precisa será la toma de decisiones.

El objetivo general de la estadística es mejorar la comprensión de hechos a partir de datos. (Moore, p. 267)

Objetivos de la estadística

- Conocer la realidad de una observación o fenómeno.
- Determinar lo típico o normal de una situación.
- Determinar los cambios que representa el fenómeno.
- Relacionar dos o más fenómenos.
- Determinar las causas que originan el fenómeno.
- Hacer estimaciones sobre el comportamiento futuro del fenómeno.
- Obtener conclusiones de un grupo menor (muestra) para hacerlas extensivas a un grupo mayor (población).
- Determinar el grado de validez y confiabilidad, ya sea de las predicciones o las conclusiones obtenidas a partir de muestras.

Divisiones de la estadística.

La estadística se divide en dos ramas:



Estadística descriptiva: Se encarga de la organización y resumen de los datos, utilizando formulas y procedimientos para su presentación tabular o gráfica. La finalidad de la estadística descriptiva es resumir la información de conjuntos más o menos numerosos de datos.

Estadística inferencial: Estudia la probabilidad de éxito de las diferentes soluciones posibles a un problema, permite la emisión de juicios o conclusiones basadas en los conocimientos que se tienen de la población o muestra.

La estadística inferencial tiene dos objetivos básicos:

1. Obtener conclusiones válidas acerca de una población sobre la base de una muestra.
2. Poder medir el grado de incertidumbre presente en dichas inferencias en términos de probabilidad.

Elementos estadísticos

Población

Conjunto de individuos o elementos de interés; la población es relativa al tipo de estudio que realizamos, o sea, la población es cuando se considera a TODOS LOS ELEMENTOS.

Muestra

Es un subconjunto de la población, se recurre a ella cuando estudiar la población tiene un alto costo o no se puede acceder a ella. Uno de los problemas en la elección de la muestra es que esta sea representativa y que permita una visión adecuada de la población, si la muestra no es representativa se podría tener conclusiones erróneas sobre la población. Ejemplo de ello es cuando nosotros deseamos saber la talla promedio de uniformes de un plantel, si el muestreo lo tomamos solo de un grupo lo más seguro es que esa muestra no sea representativa de todo el plantel (población).

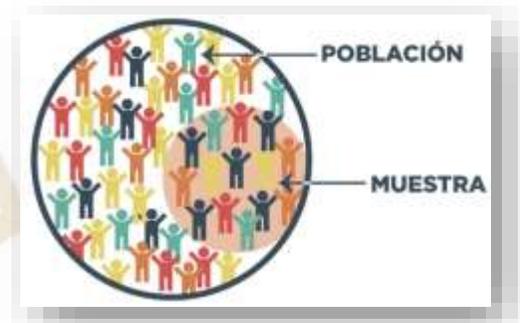


Imagen tomada de

<https://blogs.ugto.mx/enfermeriaonline/unidad-didactica-5-tamano-de-muestra/> en mayo 2021

Frecuencia

Número de veces que se repite un dato

Dato

Conocido también como información, es el valor de la variable asociada a un elemento de una población o una muestra.

Tipos de datos

Cuando realizamos recolección de información nos encontramos con una serie de datos que podemos clasificar como *cuantitativos* o *cualitativos*. Los primeros se refieren a la cantidad, son números que representan un conteo de datos (número de años, estatura o masa de una persona). Los segundos se refieren a una cualidad, categoría o atributo de los datos (genero de personas, color de inmuebles, etc.).

Variable

Es una característica de la población o muestra cuya medida puede cambiar de valor, se representa simbólicamente mediante las letras del alfabeto y según su naturaleza pueden ser cuantitativa y cualitativa.

Variable y sus tipos

Una variable estadística es una característica de una muestra o población de datos que puede adoptar diferentes valores.

Tipos de variables

Cualitativos

Discretos:

Cuando los datos pueden contarse

Ejemplo: edad, peso, número de alumnos en un grupo.

Continuos:

Datos que pueden tomar un número infinito de valores cubriendo un intervalo

Ejemplo: tiempo en llegar a la escuela, tiempo en terminar un proyecto, duración en minutos de una llamada.

Cuantitativos

Nominal:

Nombres etiquetados o categorías que pueden ordenarse

Ejemplo: Alto, medio, bajo.

Ordinal:

Cuando los datos pueden colocarse en un orden pero no es posible diferenciar entre los valores

Ejemplo: En una escuela, director, subdirector, docente, administrativo, jefe de oficina, empleado de aseo, alumnos.

Técnicas de recolección de datos

Para llevar a cabo un trabajo de investigación el investigador cuenta con gran variedad de técnicas para diseñar un plan de recolección de datos. Tales técnicas varían de acuerdo con cuatro dimensiones importantes: estructura, confiabilidad, injerencia del investigador y objetividad.



Imagen tomada de <https://online-tesis.com/tecnicas-de-recoleccion-de-datos-para-realizar-un-trabajo-de-investigacion/> en mayo de 2021

Una vez seleccionada adecuadamente la población de estudio o, en su caso, la muestra representativa, es conveniente definir la técnica o técnicas de recolección de datos y el instrumento o herramienta que nos ayudará para recabar la información requerida.

Principales técnicas de recolección de datos

- Encuesta
- Entrevista
- Observación
- Experimentación y
- Documental

Encuesta

La encuesta es una de las técnicas usada para recopilar información donde el investigador interroga a los investigados mediante preguntas a los investigados para recabar datos del comportamiento o conducta de un sujeto, ya sea de forma individual o en grupo, en situaciones reales. Imagen

Los instrumentos usados para su recolección son los cuestionarios de preguntas cerradas o abiertas, con características de validez y confiabilidad.

Entrevista

La entrevista también es una técnica de recolección de datos mediante la interacción verbal directa entre los individuos sujetos a estudio y el entrevistador, donde el entrevistador es quien formula las preguntas, a partir de un cuestionario previamente preparado, y el entrevistado responde profundizando a medida que la charla avanza. Los instrumentos usados con cuestionarios que pueden llevarse a cabo con preguntas estructuras, semiestructuradas o no estructuradas.

Observación

Es una técnica de recolección de datos que se basa en la anotación y registro de acciones, reacciones y de manera general en el comportamiento observado en los sujetos bajo estudio. El ambiente de los individuos, no se altera o modifica mientras se realiza la investigación. Las formas mediante las cuales se obtiene la información son las observaciones: heurística, de verificación, participante, de campo y experimental.

Experimentación

Técnica de recolección de datos consiste, en provocar las situaciones o eventos que darán pie al análisis del comportamiento de los sujetos al contrario de la observación, con esta técnica se recrea el ambiente en que los individuos se desenvuelven. Entre los instrumentos de recolección de datos esta la escala de actitudes y opiniones de Likert.

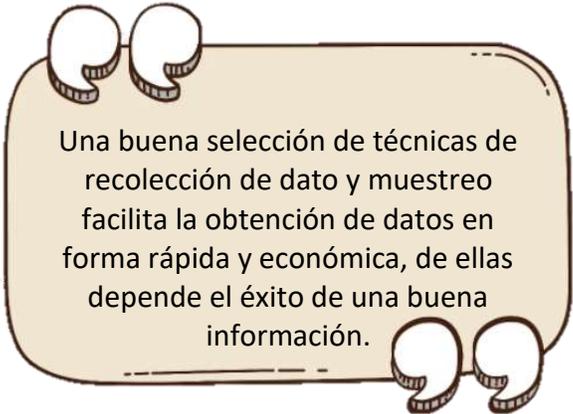
Documental

La técnica documental se basa en documentar y organizar los archivos electrónicos, periódicos, libros, revistas y cualquier medio, para revisar los resultados y las conclusiones posteriormente junto con la persona observada. Los instrumentos usados para su recolección son: la escala de apreciación (numérica, gráfica o descriptiva), investigación documental o diagrama de flujo.

Técnicas de muestreo

En ocasiones en que no es posible o conveniente realizar un censo (analizar a todos los elementos de una población), se selecciona una muestra, entendiendo como tal una parte representativa de la población.

El muestreo es, por consiguiente, un conjunto de técnicas empleadas para seleccionar la mejor muestra posible, es decir, la que se considera más representativa de la población, con la finalidad de que cuando se tengan los resultados acerca de las características que se desean estudiar, hacer una inferencia respecto de toda la población.



Una buena selección de técnicas de recolección de dato y muestreo facilita la obtención de datos en forma rápida y económica, de ellas depende el éxito de una buena información.

Tipos de muestreo

Existen diferentes criterios de clasificación de los tipos de muestreo, aunque en general pueden dividirse en dos grandes grupos: muestreo probabilístico y muestreo no probabilístico.

Muestreo Probabilístico

- *Muestreo Aleatorio Simple*
- *Muestreo Aleatorio Estratificado*
- *Muestreo Aleatorio Sistemático*
- *Muestreo Aleatorio Por Conglomerado*

Muestreo aleatorio simple

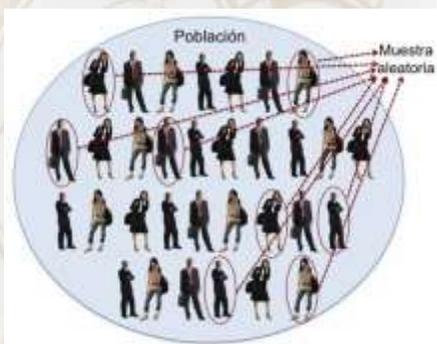


Imagen recuperada de

https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0717-95022017000100037 en mayo 2021

El muestreo aleatorio simple es la técnica de elección de una muestra que consiste en seleccionar n elementos de un total de N que conforman la población, de tal forma que todos los elementos tengan igual probabilidad de ser elegidos, (imagen de la izquierda)

Para realizar el muestreo aleatorio simple es importante contar con un listado numerado y completo (marco muestral) de todos los elementos que conforman la población y seleccionar de forma aleatoria cada uno de los elementos que integrarán la muestra. Cada elemento tiene la misma probabilidad de ser seleccionado para el estudio. Para seleccionar el número de elementos de la población puedes recurrir al método de lotería, una tabla

de números aleatorios y los números generados de forma aleatoria mediante un programa de computadora, es decir, al azar.

Pasos a seguir: determinar el tamaño de la muestra, numerar los individuos de 1 a n y elegir los elementos aleatoriamente que integran la muestra.

Ejemplo: Una empresa tiene 150 empleados. Se quiere extraer una muestra de 30 de ellos.

- Enumera a los empleados del 1 al 150
- Sortea 30 números entre los 150 trabajadores
- La muestra estará formada por los 30 empleados que salieron seleccionados de los números obtenidos.

Muestreo sistemático

El muestreo sistemático es la técnica de elección de una muestra que consiste en seleccionar n elementos de un total de N que conforman la población, de tal manera que cada elemento seleccionado de la muestra será elegido sistemáticamente conforme el siguiente procedimiento:

- a) Calcular el intervalo de selección de elementos (k) al dividir el número total de la población entre el tamaño de la muestra $k = \frac{N}{n}$
- b) Determinar un número de manera aleatoria entre 1 y k , al que llamaremos número aleatorio inicial (i).
- c) Elegir los elementos que serán seleccionados en la muestra. En general, serán los siguientes:

$$i, i + k, i + 2k, \dots, i + (n-1)k$$

Por ejemplo, el investigador tiene una población total de 100 individuos y necesita 12 sujetos. Primero elige su número de partida, 5.

Luego, el investigador elige su intervalo, $k = 100/12$ que resulta ser 8. Los miembros de su muestra serán los individuos 5, 13, 21, 29, 37, 45, 53, 61, 69, 77, 85, 93.

Muestreo estratificado

Es la técnica de elección de una muestra que consiste en subdividir en subpoblaciones o estratos a la población con el objetivo de que cada subgrupo (homogéneo entre sí en alguna característica en común) sea representativo de toda la población. Una vez determinados los estratos se saca una muestra de cada uno, la obtención se realiza independientemente en estratos diferentes de tal manera que la suma de las submuestras será la muestra deseada.

De esta manera, si una población de tamaño N se divide en R estratos, entonces la suma de los elementos de los estratos, denotados como N_i , da como resultado la población bajo estudio.

$$N = N_1 + N_2 + \dots + N_R$$

Asimismo, la muestra n será resultado de la suma de las muestras, denotadas como n_i de los diferentes estratos:

$$n = n_1 + n_2 + \dots + n_R$$

Generalmente, el tamaño de la muestra de cada estrato es proporcional al tamaño de la muestra de la población y dichas muestras se obtienen de manera aleatoria.

El procedimiento para realizar un muestreo estratificado es el siguiente:

- Definir la cantidad de n estratos.
- Verificar que la suma de los elementos de cada estrato sea del mismo tamaño de la población sin estratificar.
- Asignar a cada estrato, de acuerdo con el número de sus elementos, la parte proporcional a la muestra requerida, de manera que la suma de las muestras a extraer de cada estrato sea igual al tamaño de la muestra.
- Seleccionar, mediante muestreo aleatorio simple, los elementos de la muestra de cada uno de los estratos.

Muestreo por conglomerados

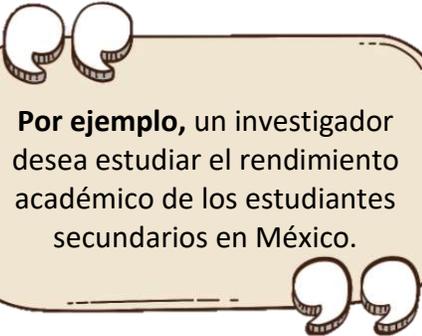
Ejemplos de la aplicación de muestreo aleatorio estratificado son: la edad, el género, el nivel socioeconómico, la religión, la nacionalidad y el nivel de estudios alcanzado.

Es la técnica de elección de una muestra que consiste en subdividir en grupos o conglomerados la población, sólo será necesario un listado (marco) de estos conglomerados y no todos los elementos, de tal modo que cada subgrupo (heterogéneo entre sí) sea representativo de la población. Se eligen al azar R conglomerados y se analizan todos los elementos que los conforman. La suma de los elementos, de cada conglomerado elegidos al azar, será igual (o muy cercano) a la muestra deseada.

Los conglomerados deben ser lo más heterogéneos posible en su conformación individual, pero ser muy similares a los otros conglomerados de la población de estudio:

El procedimiento para realizar el muestreo por conglomerados es el siguiente:

- a) Se define la población N y la muestra n .
- b) Se define el número de los R de conglomerados.
- c) Se elige de manera aleatoria un conglomerado. Si todos los elementos que lo conforman satisfacen la muestra requerida, se procede a analizarlos todos. Se seleccionan al azar tantos conglomerados como la suma de todos los elementos satisfagan la muestra requerida.



Por ejemplo, un investigador desea estudiar el rendimiento académico de los estudiantes secundarios en México.

Esta técnica de muestreo es más utilizada en la investigación de conglomerados geográficos.

1. Se puede dividir a toda la población (población de México) en diferentes conglomerados (ciudades).
2. Luego, el investigador selecciona una serie de conglomerados en función de su investigación, a través de un muestreo aleatorio simple o sistemático.
3. posteriormente, de los conglomerados seleccionados (ciudades seleccionadas al azar) el investigador puede incluir a todos los estudiantes de secundaria como sujetos o seleccionar un número de sujetos de cada conglomerado a través de un muestreo aleatorio simple o sistemático.

Muestreo No Probabilístico

- *Muestreo por conveniencia*
- *Muestreo "bola de nieve"*
- *Muestreo por cuotas*

El muestreo no probabilístico utiliza técnicas en los que no interviene el azar y, por lo tanto, se desconoce la probabilidad asociada a cada individuo para formar parte de la muestra. Normalmente estas técnicas se utilizan en estudios exploratorios o intencionales, en los cuales no es necesario proyectar los resultados. El inconveniente de este método es que no puede asegurarse la que la muestra sea representativa de la población. Los comunes son por conveniencia y bola de nieve.

A diferencia en el muestreo probabilístico, donde cada miembro de la población tiene una posibilidad conocida de ser seleccionado, en el muestreo no probabilístico, no todos los miembros de la población tienen la oportunidad de participar en el estudio.

Muestreo por conveniencia

También llamado en ocasiones “accidental” es una técnica comúnmente usada consistente en seleccionar una muestra de la población por el hecho de que sea accesible. Es decir, los individuos empleados en la investigación se seleccionan porque están fácilmente disponibles y porque sabemos que pertenecen a la población de interés, no porque hayan sido seleccionados mediante un criterio estadístico. Esta conveniencia, que se suele traducir en una gran facilidad operativa y en bajos costes de muestreo, tiene como consecuencia la imposibilidad de hacer afirmaciones generales con rigor estadístico sobre la población.



Imagen tomada de <https://www.netquest.com/blog/es/blog/es/muestreo-por-conveniencia> en mayo 2021



Imagen tomada de <https://blogs.ugto.mx/enfermeriaenlinea/unidad-didactica-5-tamano-de-muestra/> en mayo 2021.

Por ejemplo

Supongamos que queremos conocer la opinión de los estudiantes universitarios de Tabasco de la política. Una muestra probabilística requeriría acceder a un censo del total de estudiantes de todas las universidades de Tabasco con el fin de seleccionar al azar un grupo de estudiantes y encuestarlos. *Una muestra por conveniencia podría consistir en dirigirnos a 3 universidades cercanas, simplemente porque están en la población en la que reside el encuestador, y encuestar a unos cuantos individuos que acepten participar al salir de las aulas.*

Muestreo bola de nieve

El muestreo por bola de nieve es una técnica de muestreo no probabilística en la que los elementos seleccionados para ser estudiados reclutan a nuevos participantes entre sus conocidos. El nombre de "bola de nieve" proviene



Imagen tomada de <https://blogs.ugto.mx/enfermeriaenlinea/unidad-didactica-5-tamano-de-muestra/> en mayo 2021.

El muestreo no probabilístico se utiliza donde no es posible extraer un muestreo de probabilidad aleatorio *debido a consideraciones de tiempo o*

justamente de esta idea: del mismo modo que una bola de nieve al rodar por una ladera se va haciendo cada vez más y más grande, esta técnica permite que el tamaño de la muestra vaya creciendo a medida que los elementos seleccionados invitan a participar a sus conocidos.

Se usa con frecuencia para acceder a **poblaciones de baja incidencia y a individuos de difícil acceso** por parte del investigador. En proyectos en los que se quiere estudiar a un colectivo muy específico (por ejemplo, personas aficionadas a coleccionar monedas), puede resultar mucho mejor obtener una muestra a través de conocidos y amigos de los propios coleccionistas, que, mediante una selección puramente aleatoria, en la que una gran cantidad de individuos candidatos a participar serían descartados. Dado que es muy probable que un coleccionista de monedas conozca a otros coleccionistas de monedas, lo que hace de esta técnica una forma efectiva de muestrear un colectivo que de otra manera resultaría de difícil acceso para el investigador.

Muestreo por cuotas

El muestreo por cuotas es un método de muestreo no probabilístico en el que los investigadores pueden formar una muestra que involucre a individuos que representan a una población y que se eligen de acuerdo con sus rasgos o cualidades.



Imagen tomada de <https://www.questionpro.com/blog/es/muestreo-por-cuotas/#:~:text=El%20muestreo%20por%20cuotas%20es,con%20sus%20rasgos%20o%20cualidades> en mayo 2021

Los investigadores pueden decidir el rasgo según el cual se llevará a cabo la selección del subconjunto de la muestra para que ésta pueda ser efectiva en la recolección de datos y que puedan generalizar a toda la población. El subconjunto final se decidirá sólo de acuerdo con el conocimiento de la población por parte del entrevistador o investigador.

En las etapas iniciales de un estudio, los investigadores pueden recolectar datos representativos de una muestra formada

utilizando el método de muestreo por cuotas. Es muy similar al muestreo estratificado, que es un método de muestreo probabilístico. La principal diferencia entre estas dos técnicas es que, en el muestreo por cuotas, los elementos de la muestra no se seleccionan aleatoriamente de cada estrato como se hace en el muestreo aleatorio estratificado.

Por lo general, los investigadores prefieren técnicas de muestreo no probabilísticas, como el muestreo por conveniencia y el muestreo por cuotas, en situaciones en las que existen restricciones financieras o de tiempo para la investigación. Además, en algunos casos en los que la velocidad de la investigación es más valiosa que la precisión de los resultados obtenidos, se recurre a este método de muestreo.



Imagen tomada de <https://blogs.ugto.mx/enfermeriaenlinea/unidad-didactica-5-tamano-de-muestra/> en mayo 2021.

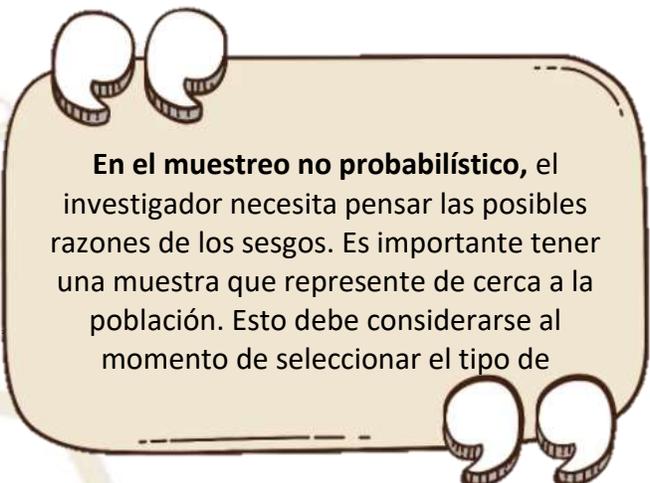
Por ejemplo

Hipotéticamente, supongamos que un investigador desea estudiar los objetivos profesionales de los empleados de una organización. En esta organización trabajan 500 empleados y estos son conocidos en conjunto como "población".

Para estudiar los objetivos de más de 500 empleados, técnicamente la muestra seleccionada debe tener un número proporcional de hombres y mujeres. Lo que significa que debe haber 250 hombres y 250 mujeres. Como esto es improbable, los grupos o estratos se seleccionan mediante el muestreo por cuotas.

¿Cuándo usar el muestreo no probabilístico?

- Este tipo de muestreo se usa para indicar si existe un rasgo o característica particular en una población.
- Esta técnica de muestreo es ampliamente utilizada cuando los investigadores realizan investigaciones cualitativas, estudios piloto o investigación exploratoria.
- El muestreo no probabilístico se usa cuando los investigadores tienen un tiempo limitado para llevar a cabo la investigación o tienen limitaciones presupuestarias.
- El muestreo no probabilístico se realiza para observar si un tema en particular necesita un análisis en profundidad.



En el muestreo no probabilístico, el investigador necesita pensar las posibles razones de los sesgos. Es importante tener una muestra que represente de cerca a la población. Esto debe considerarse al momento de seleccionar el tipo de



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

Competencias a desarrollar		TAREA No. 1	CUADRO COMPARATIVO: Técnicas de recolección de datos y muestreo
CG. 3.2 CG. 5.2 CG. 8.1	CDBM. 5 CDBM. 7		INSTRUCCIONES: Con ayuda de tu profesor y formados en binas, realizar un cuadro comparativo sobre las técnicas de recolección de datos y muestreo.

Técnicas de recolección de datos y muestreo	Definición	Características	Ventajas	Desventajas	Ejemplos de aplicación de situaciones reales.
Encuesta					
Entrevista					
Observación					



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

Experimentación					
Documental					
Muestreo probabilístico					
Muestreo no probabilístico					

PYEI-B2-LC01 Lista de Cotejo para evaluar Tarea 01 Cuadro comparativo

Asignatura:	Probabilidad y Estadística I	Bloque: I	Elementos estadísticos	Fecha:		
Nombres				Grupo:		
				Turno		
Ejercicios: Frecuencias para datos agrupados y no agrupados						
Aprendizajes Esperados			Contenidos Específico			
<p>Aplica crítica y reflexivamente los elementos estadísticos, así como la información recolectada a través de las diferentes técnicas, su tipo de variabilidad y su comportamiento en diferentes contextos.</p>			<ul style="list-style-type: none"> ▪ Elementos estadísticos ▪ Técnicas de recolección de datos ▪ Técnicas de muestreo 			
CRITERIOS				%	CUMPLE	Puntaje
					SI	NO
1. Entrega su producto terminado en el tiempo establecido por el facilitador				10%		
2. Distingue las técnicas de recolección de datos y muestreo				20%		
3. Reconoce las características de las técnicas de recolección de datos y muestreo				20%		
4. Identifica las ventajas y desventajas de las técnicas de recolección de datos y muestreo				20%		
5. Presenta ejemplos de situaciones reales sobre las técnicas de recolección de datos y muestreo				20%		
6. Se relaciona colaborativamente mostrando disposición al trabajo metódico y organizado				10%		
Calificación						
Logros obtenidos			Aspectos a mejorar			
Nombre y Firma del Coevaluador			Firma del Facilitador			



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

De la información al conocimiento



Distribución de Frecuencias

La función de la estadística descriptiva es convertir los datos recopilados en información fácil de interpretar para cualquier persona; arrojando información compacta, clara y precisa.

Representa la habilidad matemática que alcanza, constituye, presenta y detalla un conjunto de datos con la finalidad de proporcionar su uso totalmente con el soporte de tablas, medidas numéricas o gráficas.



Imagen tomada de <https://www.oas.org/ext/es/desarrollo/recursos-educacion-docente/Planes-de-Clase/Details/ArtMID/2250/ArticleID/663/-Estad-stica-Descriptiva> en mayo 2021

Organización de datos

Datos no Agrupados



Se refiere al conjunto de todos los datos tal y como han sido recopilados, son todos los valores que ha tomado la variable y que se exhiben en una lista sin haber sido organizados previamente. En caso de analizar una muestra de 100 personas, los datos no agrupados son los 100 resultados que se obtienen de este análisis.

Estos datos pueden estar presentados en su forma de aparición en una tabla de datos donde cada valor se representa de forma individual. Por lo general este conjunto comprende una cantidad de elementos menor a 30 ($n < 30$) con poca o nula repetición. El manejo de estos datos es simple, se recolectan los datos de la población de estudio y dichos datos se distribuyen en una tabla de datos y se analizan sin necesidad de formar clases con ellos

Datos Agrupados

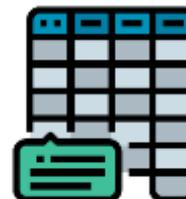
Se le llama así al resultado de organizar los datos **no agrupados**. En este caso los datos se agrupan y se ponen en un esquema que es más fácil de leer e interpretar, el cual contiene información útil adicional calculada a partir de los datos originales.

Una cantidad dada de datos que puede clasificarse ya sea por sus cualidades cualitativas o cuantitativas, y por tal agruparse para su análisis. Por lo cual pueden organizarse o clasificarse de dos formas: datos agrupados en frecuencia o en intervalos.



Dependiendo de la cantidad de datos que debamos analizar convertiremos los datos No Agrupados en Agrupados. Los listados extensos no son fáciles de entender por lo que se deben interpretar en una tabla para hacer más fácil su interpretación, es decir en las tablas de distribución de frecuencias.

Tabla de distribución de frecuencias para datos agrupados



El objetivo de las **tablas de distribución de frecuencia o tabla con datos agrupados** es de informar sobre un tema en particular, que se emplea si las variables toman un número grande de valores o la variable es continua. Se agrupan los valores en intervalos que tengan la misma amplitud denominados clases; a cada clase se le asigna su frecuencia correspondiente como se muestra a continuación:

TÍTULO							
Intervalo de clase							
Límite Inferior	Límite Superior	Marca de clase (MC)	Conteo	Frecuencia absoluta (fi)	Frecuencia absoluta acumulada (Fi)	Frecuencia relativa % (ni)	Frecuencia Relativa acumulada (Ni)
						Σ	

Dónde:

- **Título:** Nombre o enunciado que encabeza a la distribución de frecuencias
- **Intervalo o clase:** Grupos o categorías en que pueden clasificarse los datos recopilados
- **Límites inferior y superior:** Son los extremos de cada intervalo
- **Marca de clase (Mc):** Promedio o el punto medio entre los límites de cada intervalo. Se obtiene al dividir entre dos, la suma de los límites inferior y superior de cada intervalo.
- **Conteo:** Consiste en acomodar cada dato en la clase o en el intervalo que contiene su valor, es decir “contamos o localizamos” cada dato en la lista de datos NO Agrupados y lo acomodamos en la distribución de frecuencias colocando un símbolo como una “X” “|” o “✓”.
- **Frecuencia (f) o frecuencia absoluta:** Es el número de datos que pertenece a cada clase o intervalo. Se representa como f_i , donde la «i» corresponde al número de dato. Se obtiene contando las veces que aparece el dato en el conjunto de datos. La suma de las frecuencias absolutas corresponde al número total de datos, representado por la letra N :

$$\sum_{i=1}^{i=n} f_i = f_1 + f_2 + \dots + f_n = N$$

- **Frecuencia relativa (%):** Es la proporción que representan los datos de cada clase con respecto al total de datos, con un valor entre 0 y 1 y se puede expresar en tanto por ciento si lo multiplicamos por 100, agregándole el signo de % al resultado. Se representa como « n_i », siendo « i » el número de dato, y se calcula dividiendo la frecuencia absoluta de cada dato entre el número total de datos:

$$n_i = \frac{f_i}{N}$$

- **Frecuencia absoluta acumulada (F_i):** La frecuencia acumulada es la suma de las frecuencias igual o menores a la que se trata. Se obtiene sumando su frecuencia absoluta a las frecuencias absolutas de los datos que son menores que él, de este modo, la frecuencia absoluta acumulada del primer dato coincide con su frecuencia absoluta y la frecuencia absoluta acumulada del último dato coincide con el número total de datos.

$$\sum_{i=1}^{i=n} n_i = n_1 + n_2 + \dots + n_n = 1$$

- **Frecuencia relativa acumulada (N_i):** corresponde al mismo concepto que la absoluta acumulada, sin embargo, corresponde a la suma de la frecuencia relativa de un dato más la frecuencia relativa del dato anterior. Así que, la frecuencia relativa acumulada del primer dato coincide con su frecuencia relativa y la frecuencia relativa acumulada es igual a 1.

$$N_i = \frac{F_i}{N}$$

- Σ : Representa la suma de todos los valores obtenidos, si existe una diferencia entre la suma de frecuencias y el total de datos significa que no contamos o clasificamos cada dato correctamente, siendo los datos incorrectos. Al seleccionar este símbolo en una hoja de cálculo se obtiene una suma automática.
- **Análisis de la tabla de distribución de frecuencias:** Es importante explicar los datos obtenidos en la tabla de distribución de frecuencias, escribiendo afirmaciones descriptivas al respecto.

Por lo tanto, la **tabla de distribución de frecuencias o distribución de frecuencias** es un arreglo en forma de tabla que asocia cada valor de una variable con el número de veces que se presenta (**frecuencia**), la cual permite ordenar y agrupar los datos en clases, cada clase es un intervalo donde los extremos son llamados **Límites del intervalo**. Los datos pueden presentarse en forma individual o formando grupos, la segunda forma se utiliza cuando el volumen de los datos es muy grande.

YOUTUBE:

Tabla de Frecuencias para
Datos no Agrupados Ejercicios
Resueltos



<https://www.youtube.com/watch?v=iPEt789ewVM>

Determinación de clases para datos no agrupados

Distribución de datos no agrupados con variable cualitativa

Ejemplo 1: En la cooperativa de la escuela se desea saber cuáles son los tipos de aguas frescas que prefieren un grupo de estudiantes.

Bebidas preferidas				
Horchata	Limonada	Jamaica	Agua de naranja	Agua de naranja
Limonada	Jamaica	Horchata	Limonada	Jamaica
Agua de naranja	Horchata	Limonada	Jamaica	Horchata
Limonada	Agua de naranja	Jamaica	Agua de naranja	Agua de naranja

Tabla de Frecuencias Bebidas preferidas					
Clases	Conteo	Frecuencia absoluta (f_i)	Frecuencia absoluta acumulada (F_i)	Frecuencia relativa % (n_i)	Frecuencia Relativa acumulada (N_i)
Horchata	IIII	4	4	$4/20 = 0.20$	0.20
Limonada	IIIII	5	9	0.25	0.45
Jamaica	IIIII	5	14	0.25	0.70
Agua de naranja	IIIIII	6	20	0.30	1
Total		$N = 20$		$\Sigma = 1$	

Se inicia ordenando los datos obtenidos para poder elaborar la tabla de frecuencias

Procedimiento:

- **Conteo:** Número de veces que se repite la frecuencia colocando el símbolo "I"
- **Frecuencia absoluta (f_i):** cantidad de veces que se repite una frecuencia numéricamente, ejemplo: para la clase Horchata el conteo es IIII: por lo tanto, la frecuencia es 4
- **Frecuencia absoluta acumulada (F_i):** se suman las frecuencias absolutas, ejemplo para la clase Horchata se inicia con el número 4, para limonada se suma la frecuencia absoluta de Horchata más limonada: $4 + 5 = 9$
- **Frecuencia relativa % (n_i):** se obtiene de dividir la frecuencia absoluta, entre el total de frecuencia absoluta, ejemplo: para la clase Horchata: $4/20 = 0.20$, para limonada: $5/20 = 0.25$

- **Frecuencia Relativa acumulada (Ni):** Resulta de sumar la Frecuencia relativa % (ni) más la Frecuencia Relativa acumulada (Ni), ejemplo para la clase Horchata $0.20 + 0.25 = 0.45$

Análisis: Se obtiene como resultado que la vista preferida de los estudiantes es el Agua de naranja y la que menos prefieren es el agua de horchata, con un porcentaje de 10% de diferencia entre una y otra. El agua de limonada y Jamaica tienen un porcentaje de 25% de preferencia, con un 5% menos respecto al agua de naranja.

Distribución de Datos no agrupados con variable cuantitativa

Ejemplo 2: Realizar la tabla de distribución del salario que reciben 56 empleados de una empresa que se dedica al reciclado de basura.

Salario de los empleados en miles							
28	23	19	19	18	23	20	20
27	21	23	17	15	16	19	19
17	31	18	17	25	13	20	17
24	26	22	26	11	19	15	23
24	20	24	21	22	21	19	25
17	23	20	18	27	24	20	29
22	21	12	21	26	16	16	17

Tabla de Frecuencia Salarios											
Clase	Intervalo de clase		Límites reales		Marca de clase (m)	Frecuencia Absoluta (fi)		Frecuencia Acumulada (Fi)		Frecuencia complementaria	
	Límite inferior	Límite superior	Inferior	Superior		Simple	Relativa	Simple	Relativa	Simple	Relativa
1	11	14	10.5	14.5	12.5	2	0.035	2	0.035	54	0.96
2	15	18	14.5	18.5	16.5	14	0.25	16	0.29	40	0.71
3	19	22	18.5	22.5	20.5	21	0.375	37	0.66	19	0.34
4	23	26	22.5	26.5	24.5	14	0.25	51	0.91	5	0.09
5	27	31	26.5	31.5	29.0	5	0.09	56	1	0	0
					$\Sigma =$	56	1				

Procedimiento:

- El **rango** se obtiene a partir del valor más grande y pequeño:
Rango: $31 - 11 = 20$
- **Intervalo de clase:** Decidir el número de intervalos y la cantidad de elementos depende únicamente de quien diseña la tabla. En ocasiones se toma el criterio de la raíz cuadrada, raíz cúbica o la regla de Storges que está dada por la ecuación $NC = 1 + 3.3 \log N$ (N número de datos u observaciones). Por lo tanto, se sugiere:
 - ✓ Los intervalos no deben ser ni muchos ni pocos, teniendo el mismo ancho
 - ✓ Conviene calcular la raíz cuadrada de N (Diferencia entre el límite más pequeño y el más grande)
- Para obtener la amplitud del intervalo se divide el rango entre el número de clases que se seleccionaron, en caso de ser fraccionario se debe fraccionar a la unidad de variación siguiente:

$$A = \frac{\text{Rango (R)}}{\text{Número de intervalos (K)}} = \frac{20}{5} = 4$$

- **Intervalos de clase o intervalo exacto** se forman por el límite inferior y superior.
El límite inferior de la primera clase es el valor del menor de los datos: $11 + 4$
Para el límite superior se resta una unidad de variación al límite inferior anterior de la siguiente clase: $11 + 4 - 1$.
El valor 31 no se encuentra en la tabla, como el último intervalo de clases termina en 31 se realiza un ajuste, cambiando el 30 por el 31, con lo que todos los datos estarán incluidos

- **Límites reales o exactos:** se obtienen restando media unidad de variación al límite inferior y al límite superior se le suma: al ser la unida de variación 1000, media unidad es 500, representándolo en la tabla como 0.5
- **Marca de clase:** punto medio entre extremos de cada intervalo. Se suman los extremos del intervalo de clase y se divide entre 2

$$\begin{array}{rclclcl} \text{Límites:} & 11 & + & 14 & = & 25 \\ \text{Límites reales} & 10.5 & + & 14.5 & = & 25 \end{array}$$

Dividimos entre 2 y se obtiene 12.5 que es la marca de clase. Para obtener las marcas de clase siguientes es sumar a la primera el tamaño del intervalo.

- **Frecuencia absoluta:** En el intervalo de clase 11 – 14 tenemos dos datos: el 11 y el 13, anotemos en la columna 5 de la frecuencia absoluta simple (f) el número 2.
En el intervalo 19 – 22 la frecuencia absoluta simple (f) es 21.
Frecuencia absoluta simple: 2 es el número de veces que se repite la clase
Frecuencia relativa: 2 se divide entre las frecuencias $2/56 = 0.035$
- **Frecuencia acumulada:** suma de todas las frecuencias simples de las clases anteriores y la correspondiente a la clase que se está trabajando.

Frecuencia acumulada simple: En la primera clase se tienen sólo dos personas, la frecuencia acumulada es igual a 2; en la segunda clase tenemos 14 y las dos de la clase anterior suman 16, que será la frecuencia acumulada en la clase 2; en la tercera clase suman 21 que se tienen y las 16 anteriores, por lo que la frecuencia acumulada es 37, así se continúa hasta el final.

Frecuencia acumulada relativa: es el cociente de la frecuencia acumulada simple y la suma de las frecuencias simples. Para la primera clase: $2 / 56 = 0.035 = 3.5 \%$

- **Frecuencia complementaria:** *Frecuencia complementaria simple:* se obtiene restando al total de las frecuencias simples y las frecuencias simples de la primera clase (columna 5), las subsecuentes se obtienen restando la frecuencia simple y de la clase la frecuencia complementaria de la clase anterior, en la última clase su valor es cero.

Frecuencia complementaria relativa: se obtiene dividiendo la frecuencia complementaria simple (f) entre el total de frecuencia simple (6); ejemplo: en la primera clase se divide $54/56 = 0.96$



YOUTUBE:
TABLA DE FRECUENCIAS
AGRUPADA EN INTERVALOS
Super fácil para principiantes

<https://www.youtube.com/watch?v=VNMck8wco98>

Distribución de datos agrupados con variable cuantitativa

Ejemplo 3:

En una escuela de educación secundaria cuentan con una matrícula para tres grados, ya que existen un total de 1080 estudiantes inscritos, de edad entre 12 y 16 años de los cuales el porcentaje de hombres es 40 y el 60 es mujer, para el desfile de la revolución se pretende hacer filas por estaturas del más alto al más bajo. ¿Cuántas filas se van a formar? ¿Cuál es la estatura inicial y final para la primera fila? ¿Cuál es el rango de referencia para seleccionar a los estudiantes en cada fila?

Al hacer el ejercicio para obtener las alturas los resultados fueron los siguientes:

Registro de estaturas									
144	144	155	156	156	158	158	160	160	165
166	166	167	167	180	180	180	180	188	188

Para obtener los rangos debemos se identifica el valor mínimo y el valor máximo.

$$R = x_{MAX} - x_{MIN}$$

$$R = 188 - 144$$

$$R = 44$$

Para saber los intervalos se aplica la regla de Sturges:

$$k = 1 + 3,322 \text{Long } 20$$

$$k = 4,32$$

Se obtiene la amplitud:

$$A = \frac{R}{k}$$

$$A = \frac{44}{5}$$

$$A = 8.8$$

Para calcular la marca se realiza lo siguiente:

$$x = \frac{144 + 152}{2} = 148$$

Se repite en cada intervalo. A continuación, se identifica la frecuencia utilizando los datos que teníamos al inicio donde se anotaron las estaturas.

Tabla de registro de estaturas					
Intervalos /Estatura)	(Clase	Marca De Clase	Frecuencia (fi)	Frecuencia Relativa (ni)	Frecuencia Acumulada (Ni)
144 – 152		148	2	0.1	2
152 – 161		157	7	0.35	9
161-170		166	4	0.2	13
170-179		175	1	0.05	14
179- 188		184	6	0.3	20

Se calcula la frecuencia relativa:

$$fr_i = \frac{fi}{n} = \frac{2}{20} = 0.1$$

$$fr_i = \frac{fi}{n} = \frac{7}{20} = 0.35$$

$$fr_i = \frac{fi}{n} = \frac{4}{20} = 0.2$$

$$fr_i = \frac{fi}{n} = \frac{1}{20} = 0.05$$

$$fr_i = \frac{fi}{n} = \frac{6}{20} = 0.3$$

Para calcular la frecuencia acumulada se suman las frecuencias, comenzando en la segunda fila ya que el valor inicial es el número de la primera fila.

$$F = f_1 = 2$$

$$F = f_1 + f_2 = 2 + 7 = 9$$

$$F = f_1 + f_2 = 9 + 4 = 13$$

$$F = f_1 + f_2 = 13 + 1 = 14$$

$$F = f_1 + f_2 = 14 + 6 = 20$$

Determinación de clases para datos agrupados

Con frecuencia encontramos en estudios o investigaciones, variables discretas las cuales finalmente su objetivo es tomar valores distintos, ejemplos de estos casos son: el número de hijos que tienen las parejas en una determinada zona geográfica, el número tinacos vendidos por una empresa durante un mes, el número de empleados que tiene el Colegio de Bachilleres de Tabasco, el número de partidos a celebrarse por jornada en s de la liga mexicana de futbol torneo de clausura, Cuántos viajes realiza un autobús en la ruta de Villahermosa-Cárdenas Tabasco, o bien, ahora si hablamos de variables continuas se multiplican las posibilidades para encontrar ejemplos, son tantos los posibles valores de variable que prácticamente resulta inadecuado construir una distribución de frecuencias con base a cada valor. Para este tipo de situaciones, se recurre al agrupamiento de datos en intervalos para luego contabilizar el número de casos que pertenecen a cada intervalo.



Imagen tomada de: https://www.emprender-facil.com/por-que-es-importante-lo-estadistico-para-los-emprendedores/#google_vignette en mayo 2021.

Ejemplo 1:

La liga Mexicana de Fútbol de Expansión requiere la estadística de cuántas tarjetas amarillas obtuvieron los equipos durante la primera mitad del torneo, por lo cual requiere efectuar un análisis del problema. Para ello recolectó la información de las cédulas arbitrales que entregaron a la Federación los árbitros asignados para dichos encuentros, el resultado es la siguiente información que se presenta:

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

Primero se debe determinar el número de intervalos o clases necesarios para resumir esta información. Existen varias reglas para hacerlo, la más común es la regla de Herbert Sturges, cuya fórmula es la siguiente:

$$\text{Número de intervalos } K = 1 + 3.322 \text{ Log } n$$

$$\text{En este ejemplo } K = 1 + 3.322 (48)$$

$$= 1 + 3.322(1.68) = 1 + 5.58 = \mathbf{6.58}$$

valor que se redondea al entero próximo en este caso es **7** puesto que el número de intervalos siempre debe ser entero.

Enseguida se calcula el número de unidades de variación en los datos a esto se le llama Rango, simbolizado con **R** y es la diferencia entre el dato mayor y el menor.

$$\text{En este ejemplo } R = 29 - 10 = 19$$

A continuación, se requiere calcular la anchura o amplitud que deberá tener cada intervalo, la cual deberá ser la misma para cada uno y en adelante la conoceremos con la letra mayúscula **A** para obtenerla se divide el rango entre el número de intervalos.

$$\text{Amplitud de Intervalo} = \frac{R \text{ (Rango)}}{K \text{ (Número de intervalos)}} = \frac{19}{7} = 2.7$$

Es importante señalar que la amplitud tenga tantas partes decimales como las tenga la variable, si la variable tiene distintas cantidades de dígitos decimales entonces hay que homogeneizarlos para después obtener esa misma cantidad de dígitos en la amplitud, como en este ejemplo son valores enteros, entonces la amplitud deberá ser con el número entero inmediato mayor; es decir a 3; el no

seguir esta regla, existe la posibilidad de no poder acomodar datos en alguno de los intervalos; por lo que es muy importante cumplir con lo antes dispuesto.

En estadística se le conoce como unidad de variación a la medida que se está utilizando. Por ejemplo, cuando hablamos de peso nos referimos a *kg* como unidad de variación; si hablamos de estatura, la unidad de variación estará dada en centímetros, aunque posiblemente podríamos tomar como unidad de variación los milímetros, poco acostumbrado; en cambio, no sería correcto tomar kilómetros. En nuestro ejemplo la unidad de variación es la pieza.



El intervalo de clase se forma por el límite inferior y el límite superior. el límite inferior de la primera clase es el valor del menor de los datos, en nuestro ejemplo es el número 10 y como debe contener tres valores discretos estos son 10, 11 y 12, entonces el intervalo termina en 12. el siguiente intervalo inicia en 13 y termina 15, y así sucesivamente hasta llegar al séptimo intervalo de 28 a 30. Los intervalos serán registrados en la primera columna de la tabla, posterior a esto, se contabiliza el número de datos que cumplen con cada clase o intervalo, lo cual será el número que se coloca en la segunda columna y que conocemos como frecuencia absoluta, en este contexto de agrupamiento de datos. Se está en condiciones de elaborar la distribución de frecuencias absolutas para el número de tarjetas amarillas de cada equipo de la liga.

Clase	Intervalo de la Clase		Frecuencia Absoluta
	Límite Inferior	Límite Superior	
1	10	12	2
2	13	15	6
3	16	18	10
4	19	21	16
5	22	24	8
6	25	27	5
7	28	30	1
			48

Límites de los intervalos de clase para variables numéricas discretas

Los valores extremos de un intervalo de clase se les conoce como límites. Al extremo izquierdo le denominamos límite inferior y al extremo derecho del intervalo le denomina límite superior. Por ejemplo, para el primer intervalo de la distribución anterior, el número 10 es el límite inferior y el número 12, el límite superior. También existen los límites reales de cada intervalo. Para el caso de variables discretas, estos límites son teóricos y se pueden interpretar como aquellos valores que pueden ser redondeados al entero más próximo, por ejemplo: el número 10 puede redondear a cualquier número real mayor a 9.5, de igual manera el número 12 puede redondear a cualquier

número real menor que 12.5. Tanto el 9.5 como el 12.5 son los límites reales de nuestro primer intervalo.

Partiendo del análisis hecho en el párrafo anterior, se puede decir que el límite real inferior de cualquier intervalo de variable discreta se obtiene promediando el límite superior del intervalo que le antecede con el límite inferior de su intervalo; de igual manera su límite real superior se obtiene promediando su límite superior con el límite inferior del intervalo que le precede. Estos límites reales serán muy útiles posteriormente cuando se calculen medidas estadísticas descriptivas, tales como moda y mediana.

También un valor representativo de cada intervalo lo es su marca de clase o punto medio, que se representa con m_c y se calculará promediando los límites superior e inferior de cada clase.

Para el ejemplo anterior tendríamos entonces lo siguiente:

Clase	Intervalo de la Clase		Frecuencia Absoluta	Marca de Clase m_c	Límites Reales	
	Límite Inferior	Límite Superior			Límite Inferior	Límite Superior
1	10	12	2	11	9.5	12.5
2	13	15	6	14	12.5	15.5
3	16	18	10	17	15.5	18.5
4	19	21	16	20	18.5	21.5
5	22	24	8	23	21.5	24.5
6	25	27	5	26	24.5	27.5
7	28	30	1	29	27.5	30.5
			48			

Ejemplo 2.

En la ciudad de Villahermosa se desea saber el rango de precios de una suite con todas las comodidades que puede ofrecer un hotel de la ciudad, para lo cual se recolectó información de 20 hotel de la ciudad. Los siguientes datos representan cantidades en miles de pesos:



6.8	7.2	3.9	1.4	4.1	5.2	5.8	2.7	3.0	3.5
2.4	5.3	5.7	7.4	4.8	5.8	5.4	4.0	5.5	4.4

Organiza y construye una tabla de la distribución de frecuencias, incluyendo intervalos y las marcas de clase.

Ante todo, la variable de estudio es continua, por tal motivo, los límites reales coincidirán con los límites de los intervalos.

Procedimiento:

Como primer paso procedemos a calcular el número de intervalos mediante la fórmula de Sturges $K = 1 + 3.322 \text{ Log } n$

Al sustituir tenemos:

$$K = 1 + 3.322 \text{ Log } 20 = 5.32$$

Al redondear al entero próximo nos da un valor de $K = 5$

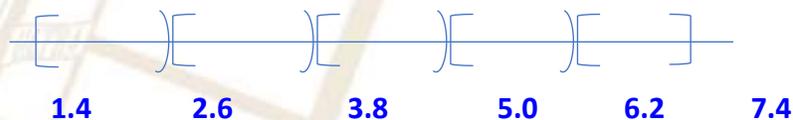
Encuentra el Rango. $R = \text{Dato mayor} - \text{dato menor.}$, en este caso $R = 7.4 - 1.4 = 6$, Por tanto, el Rango tiene un valor de 6.

Determinar la amplitud,

$$A = \frac{\text{Rango } (R)}{\text{Número de intervalos } (K)} = \frac{6}{5} = 1.2$$

En este caso no es necesario ajustar al alza el resultado puesto que los datos contienen hasta décimas, por lo tanto, nuestra amplitud queda tal y cual se obtuvo.

Procedemos a construir los intervalos de la siguiente manera: El primero de ellos inicia con el dato menor 1.4 y como la amplitud es de 1.2, la sumamos al límite inferior y obtenemos al límite superior, es decir $1.4 + 1.2 = 2.6$, este número es el límite superior del primer intervalo. Para garantizar la continuidad de la variable, se recurre a los intervalos semicerrados por la izquierda los cuales se representan con los símbolos $[Li, Ls)$ "el corchete" implica que el extremo inferior se incluye en el intervalo o que forma parte de él y si uno o más datos coinciden con él, contabilizarán para este intervalo. "El paréntesis" implica la exclusión del extremo superior del intervalo, y si uno o más datos coinciden con él, no se contabilizan en este intervalo, pero si en su intervalo sucesor.



El último intervalo debe ser cerrado por ambos lados, puesto que ya no existen más intervalos y para evitar que un valor extremo quede fuera de la contabilización. Ahora se construyen las distribuciones de frecuencias.

Clase	Intervalo de la Clase		Marca de Clase m_c	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Absoluta Acumulada	Frecuencia Relativa Acumulada
	Límite Inferior	Límite Superior					
1	[1.4	2.6)	2.0	2	0.10	2	0.10
2	[2.6	3.8)	3.2	3	0.15	5	0.25
3	[3.8	5.0)	4.4	5	0.25	10	0.50
4	[5.0	6.2)	5.6	7	0.35	17	0.85
5	[6.2	7.4]	6.8	3	0.15	20	1.00
				20			

YOUTUBE:

Tabla de Frecuencias para Datos Agrupados en Intervalos - Ejercicios Resueltos

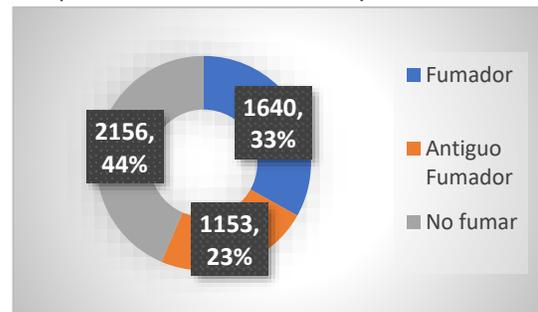


<https://www.youtube.com/watch?v=6ygaz0vECzY>

Criterios de representación y análisis de gráficas

Cuando se dispone de datos de una población, y antes de abordar análisis estadísticos más complejos, un primer paso consiste en presentar esa información de forma que ésta se pueda visualizar de una manera más sistemática y resumida. Los datos que nos interesan dependen, en cada caso, del tipo de variables que estemos manejando.

Para variables categóricas, como el sexo, profesión, etc., se quiere conocer la frecuencia y el porcentaje del total de casos que "caen" en cada categoría. Una forma muy sencilla de representar gráficamente estos resultados es mediante diagramas de barras o diagramas de sectores. En los **gráficos de sectores**, también conocidos como diagramas de "tartas", se divide un círculo en tantas porciones como clases tenga la variable, de modo que a cada clase le corresponde un arco de círculo proporcional a su frecuencia absoluta o relativa. Un ejemplo se muestra en el gráfico de la derecha.



Diagramas de "tartas"-Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.

Como se puede observar, la información que se debe mostrar en cada sector hace referencia al número de casos dentro de cada categoría y al porcentaje del total que estos representan. Si el número de categorías es excesivamente grande, la imagen proporcionada por el gráfico de sectores no es lo suficientemente clara y por lo tanto la situación ideal es cuando hay alrededor de tres categorías. En este caso se pueden apreciar con claridad dichos subgrupos.

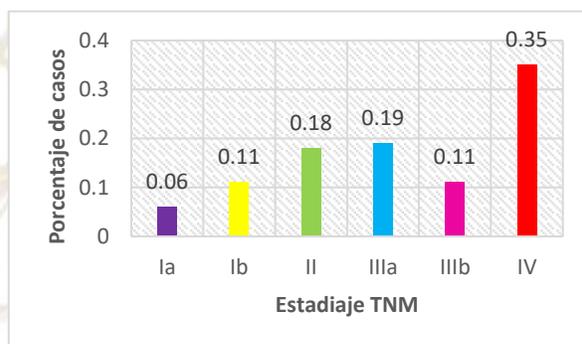
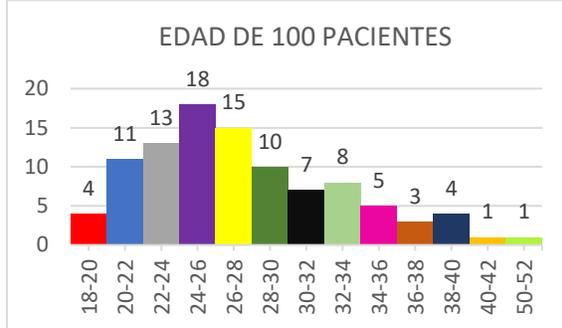


Diagrama de barras-Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.

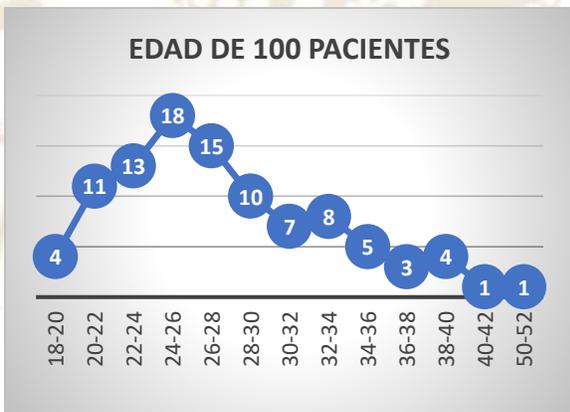
Los **diagramas de barras** son similares a los gráficos de sectores. Se representan tantas barras como categorías tiene la variable, de modo que la altura de cada una de ellas sea proporcional a la frecuencia o porcentaje de casos en cada clase (gráfico de la izquierda). Estos mismos gráficos pueden utilizarse también para describir **variables numéricas discretas** que toman pocos valores (número de hijos, número de recidivas, etc.).

Para **variables numéricas continuas**, tales como la edad, la tensión arterial o el índice de masa corporal, el tipo de gráfico más utilizado es el **histograma**. Para construir un gráfico de este tipo, se divide el rango de valores de la variable en intervalos de igual amplitud, representando sobre cada intervalo un rectángulo que tiene a este segmento como base. El criterio para calcular la altura de cada rectángulo es el de mantener la proporcionalidad entre las frecuencias absolutas (o relativas) de los datos en cada intervalo y el área de los rectángulos. Como ejemplo, La tabla de la derecha, nos indica la distribución de frecuencias de la edad de 100 pacientes, comprendida entre los 18 y 52 años. Si se divide este rango en intervalos de dos años, el primer



Histograma. Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.

tramo está comprendido entre los 18 y 19 años, entre los que se encuentra el $\frac{4}{100} = 4\%$ del total.



Polígono de frecuencia. Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.

Por lo tanto, la primera barra tendrá altura proporcional a 4. Procediendo así sucesivamente, se construye el histograma que se muestra.

Uniendo los puntos medios del extremo superior de las barras del histograma, se obtiene una imagen que se llama **polígono de frecuencias**. Dicha figura pretende mostrar, de la forma más simple, en qué rangos se encuentra la mayor parte de los datos. Un ejemplo, utilizando los datos anteriores, se presenta en la gráfica de la izquierda.

Comparación de dos o más grupos

Cuando se quieren comparar las observaciones tomadas en dos o más grupos de individuos una vez más el método estadístico a utilizar, así como los gráficos apropiados para visualizar esa relación, dependen del tipo de variables que estemos manejando. Cuando se trabaja con dos variables cualitativas podemos seguir empleando gráficos de barras o de sectores. Podemos querer determinar, por ejemplo, si en una muestra dada, la frecuencia de sujetos que padecen una enfermedad coronaria es más frecuente en aquellos que tienen algún familiar con antecedentes cardiacos.



A partir de dicha muestra podemos representar, como se hace en el gráfico siguiente. dos grupos de barras: uno para los sujetos con antecedentes cardiacos familiares y otro para los que no tienen este tipo de antecedentes. En cada grupo, se dibujan dos barras representando el porcentaje de pacientes que tienen o no alguna enfermedad coronaria. No se debe olvidar que cuando los tamaños de las dos poblaciones son diferentes, es conveniente utilizar las frecuencias relativas, ya que en otro caso el gráfico podría resultar engañoso.

Por último, hay que señalar que también en esta situación pueden utilizarse los ya conocidos gráficos de barras, representando aquí como altura de cada barra el valor medio de la variable de interés. Los gráficos de líneas pueden resultar también especialmente interesantes, sobre todo cuando interesa estudiar tendencias a lo largo del tiempo. No son más que una serie de puntos conectados entre sí mediante rectas, donde cada punto puede representar distintas cosas según lo que nos interese en cada momento (el valor medio de una variable, porcentaje de casos en una categoría, el valor máximo en cada grupo, etc.).

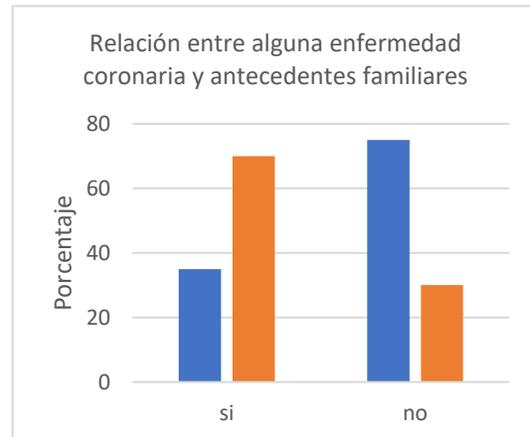


Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.

HISTOGRAMA

Toda vez que se ha hecho el análisis de frecuencias, existe en estadística, un conjunto de imágenes gráficas, las cuales, combinando distintos tipos de colores, sombreados, puntos, líneas, símbolos, números o texto, etcétera, y un sistema de referencia (coordenadas), nos permite la representación en forma más resumida y total del experimento o fenómeno en estudio. Los gráficos son muy útiles como apoyos e incluso sustitutos de las tablas o distribuciones y como una herramienta para el análisis de los datos, lo que los convierte en el medio más efectivo para la presentación, descripción, resumen y análisis de la información.

En este curso se abordarán los gráficos estadísticos como un vehículo de presentación y herramienta de análisis que permita observar tendencias presentes en los datos obtenidos al realizar un estudio.

Una manera sencilla de diferenciar los segmentos es sombreándolos de claro a oscuro, siendo el de mayor tamaño el más claro y el de menor tamaño el más oscuro.

Presentación de Datos: Después de la Organización de los datos y su presentación en Tablas Estadísticas, la información contenida en una tabla estadística también se puede

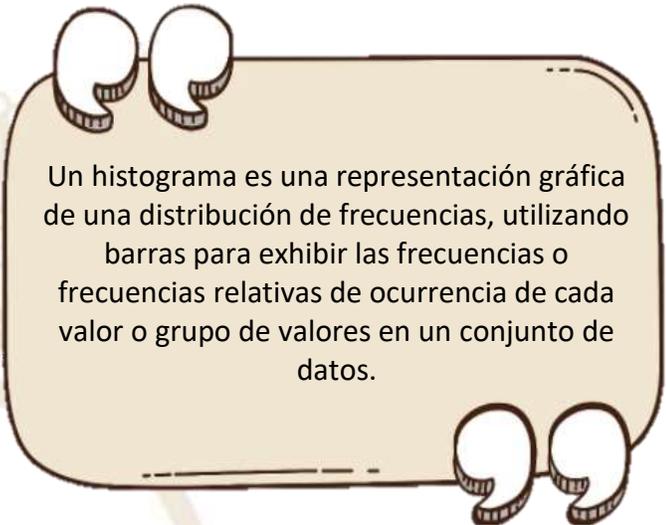


Imagen tomada de <https://soloindustriales.com/histograma/> en mayo 2021.

presentar mediante gráficas, siendo las más comunes para variables discretas (datos no agrupados) las de: Barras y circulares o de pastel; y para variables continuas (datos agrupados) el histograma, polígono de frecuencias y ojiva. Estos gráficos no son los únicos para la presentación y análisis de datos estadísticos, pero sí los más comunes y utilizados. Iniciaremos analizando el Histograma y el Polígono de Frecuencia.

Un histograma es utilizado para:

1. Resumir un conjunto de datos para una sencilla comprensión visual de sus características generales, tales como valores típicos, extensión o variación y forma.
2. Sugerir modelos de probabilidad o transformaciones para subsecuentes análisis.
3. Detectar un comportamiento inesperado o valores inusuales en los datos. Un histograma es una útil herramienta de diagnóstico para detectar valores periféricos, formas atípicas en el histograma a menudo proveen importantes pistas hacia la naturaleza del sistema o proceso que genera los datos.



Un histograma es una representación gráfica de una distribución de frecuencias, utilizando barras para exhibir las frecuencias o frecuencias relativas de ocurrencia de cada valor o grupo de valores en un conjunto de datos.

Los datos están agrupados en intervalos de la misma anchura, son mutuamente exclusivos, e incluyen todos los posibles datos. Para construir un histograma, se dibujará básicamente un diagrama de barras, sin espacios entre éstas, colocando en el eje horizontal las marcas de clase o los límites de cada intervalo en los extremos de las barras y en el eje vertical, una escala en la que se localizan las frecuencias correspondientes de cada intervalo de clase. Las barras se dibujan centradas en la marca de clase y con una altura igual a la frecuencia del intervalo.



Imagen tomada de <https://youtu.be/XFiDGydsGvU> en mayo de 2021

Ejemplo 1:

Se registran los tiempos de las llamadas recibidas en un call center, y se obtiene la siguiente tabla de frecuencias con datos agrupados. Construir un histograma de frecuencias.

Clase	Intervalo de la Clase		Marca de Clase m_c	Frecuencia Absoluta	Frecuencia Absoluta Acumulada	Frecuencia Porcentual
	Límite Inferior	Límite Superior				
1	[0	10)	5	2	2	5%
2	[10	20)	15	6	8	15%
3	[20	30)	25	12	20	30%
4	[30	40)	35	10	30	25%
5	[40	50)	45	6	36	15%
6	[50	60]	55	4	40	10%

MATEMOVIL:

Histogramas, ejemplos y ejercicios



<https://matemovil.com/histogramas-ejemplos-y-ejercicios/>



TABASCO

"Educación que genera cambio"



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

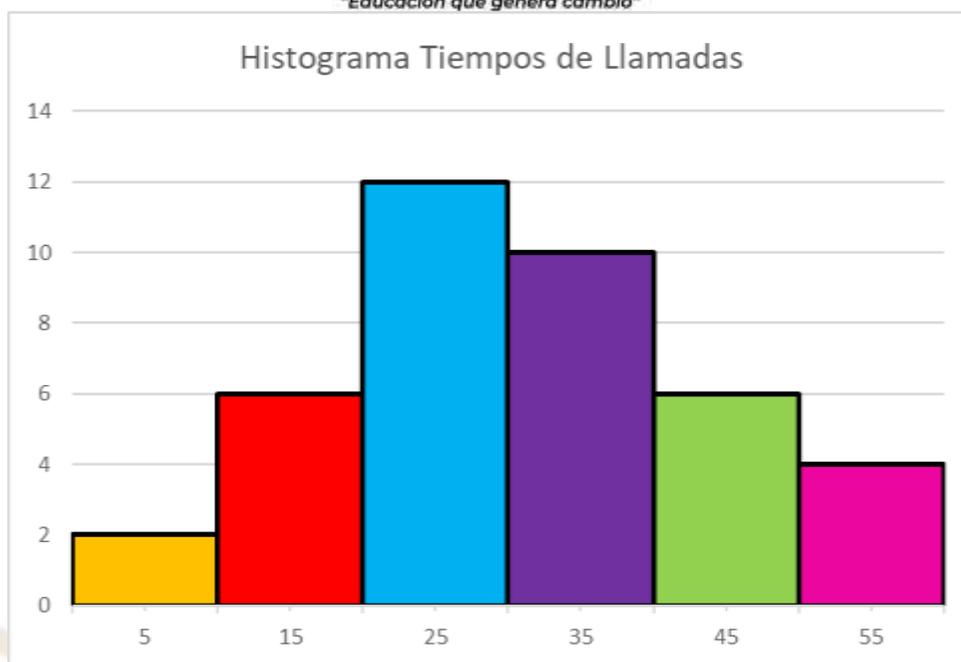


Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.

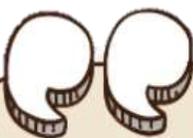
POLÍGONO DE FRECUENCIA

Un polígono de frecuencias **es una herramienta gráfica que se emplea a partir de un histograma de frecuencia** (es decir, otro tipo de gráfico que expresa las frecuencias mediante columnas verticales). Para ello, se unen con una línea los distintos puntos medios de las columnas del histograma, sin dejar *espacio* entre una y otra, logrando así una *forma geométrica* o polígono.



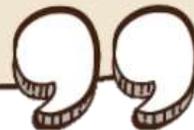
Con esta herramienta gráfica **pueden representarse variables cuantitativas o distribuciones diferentes**, cosa que tradicionalmente no hace un histograma, de un modo rápido y sencillo. Además, cuenta con la virtud de ser apreciable a simple vista.

Por esta razón es sumamente empleado dentro de las *ciencias sociales y ciencias económicas*, permitiendo así establecer *comparaciones* útiles entre los distintos resultados de un mismo *proceso*.



¿Para qué sirve un polígono de frecuencias?

Se emplean los polígonos de frecuencias cuando es necesario graficar o resaltar distintas distribuciones conjuntas o bien una clasificación cruzada de una variable cuantitativa continua, junto con otra variable cualitativa o cuantitativa discreta, todo dentro de un mismo gráfico.



Características del polígono de frecuencias

Los polígonos de frecuencia se conforman uniendo los puntos medios de cada fase o columna mediante segmentos de recta, de modo que consisten en un tipo de representación visual de la información cuantitativa. Los datos de la tabla se hallan siempre por debajo de la curvatura del polígono, y su punto más alto es siempre el de mayor frecuencia del conjunto.

Retomando los datos del ejemplo anterior nuestro Polígono de frecuencia quedaría de la siguiente manera:

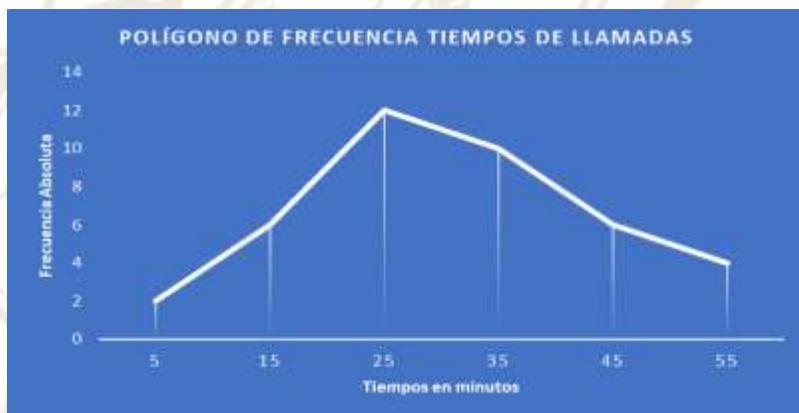


Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.

<https://youtu.be/Eoja52l-B2o>
Da clic a COBACHITO simultáneamente con la tecla CTRL



GRÁFICA CIRCULAR



Un gráfico circular es una representación en una dimensión utilizada para percibir de una forma más rápida el peso o la proporción de las categorías sobre la frecuencia total.

En otras palabras, un gráfico circular es la representación de la frecuencia relativa de las categorías de una variable, tanto cualitativa como cuantitativa. No obstante, y a pesar de lo mencionado anteriormente, debemos destacar que el gráfico circular puede representar también frecuencias absolutas.

También recibe otros nombres como **gráfico de sectores** o **gráfico de pastel**.

Es importante tener en cuenta que las variables cualitativas o que pretenden representar un orden o una categoría siempre tienen que ir ligadas a un índice numérico mayor que 0 para que pueda aparecer en el gráfico y se puedan calcular los estadísticos correspondientes.

Normalmente este tipo de gráfico va acompañado de etiquetas encima de las proporciones indicando el peso que tienen las categorías respecto al total. Si hay pocas categorías se puede prescindir de añadir las etiquetas, pero siempre es recomendable para facilitar la comprensión.

En estadística, es una herramienta útil para representar conjuntos de datos tanto discretos como continuos.

Clave del gráfico circular

Para recordar este tipo de gráfico de forma sencilla y rápida, tenemos que pensar en una pizza. De la misma forma que una pizza es redonda, el gráfico circular, tal y como dice su nombre, también lo es.

Así pues, del mismo modo que servimos una pizza, a porciones, también repartimos el gráfico circular en porciones.

Ventajas y desventajas del gráfico circular

El gráfico circular presenta unas ventajas, pero también unas desventajas.

Ventajas

- Se identifica mucho más rápido las proporciones de las categorías mediante este gráfico que empleando una tabla.

Desventajas

- Cada gráfico corresponde a una variable. Es decir, no podemos representar dos variables en un mismo gráfico. Un ejemplo sería querer representar el número de estudiantes varones que están adscritos al turno matutino y los que están adscritos al turno vespertino. Para ello tendríamos que generar dos gráficos circulares, uno para el matutino y otro para vespertino. Una opción para solo generar un gráfico sería representar los estudiantes adscritos a ambos turnos.
- Si hay muchas categorías nos puede resultar difícil diferenciar entre ellas y puede llegar al punto de no ser agradable para la vista.

Ejemplos de gráficos circulares

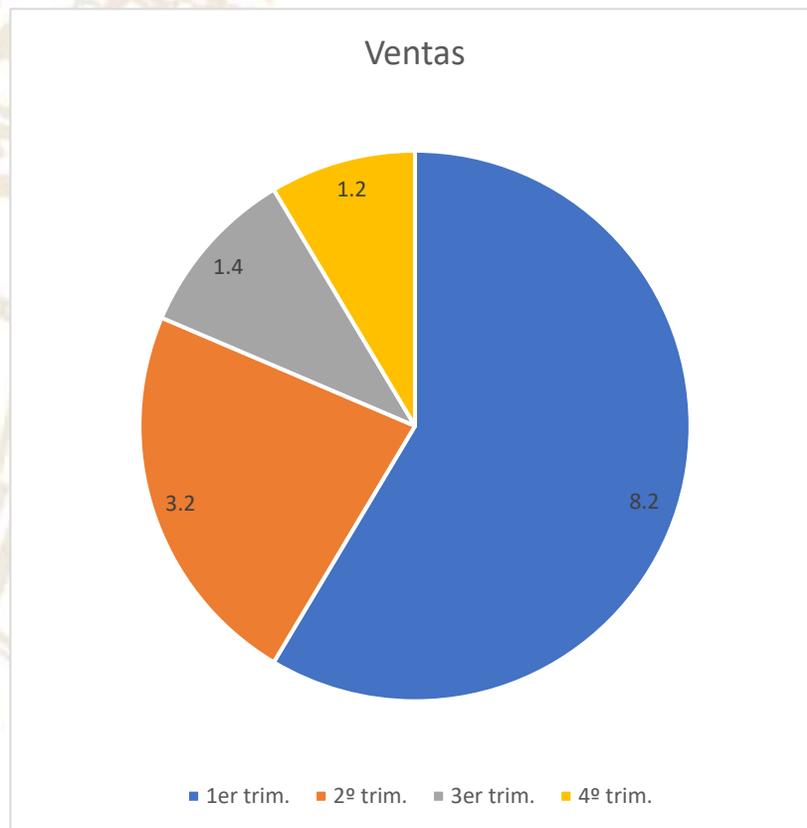


Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.



Artículos Defectuosos

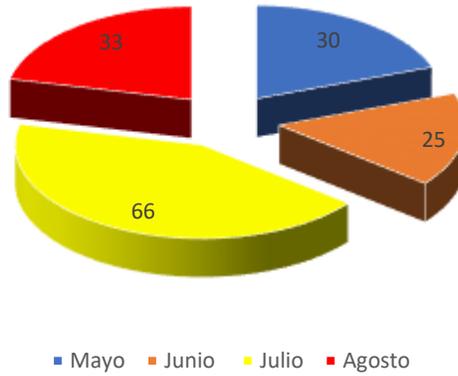


Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.

Goles anotados

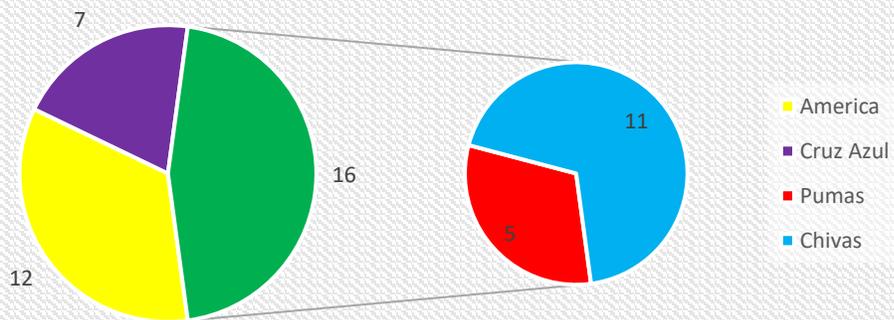


Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.



Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.

¿Como se elabora una gráfica circular?

Supongamos que levanto una encuesta a 60 personas y le pregunto cual es su sabor de pastel favorito, el resultado se muestra a continuación

Sabor	frecuencia absoluta	frecuencia relativa %	Grados
Chocolate	10		
Vainilla	16		
Fresa	13		
Queso de bola	21		
Total	60		

Iniciamos por calcular lo que sabemos hacer, es decir calcular nuestra frecuencia relativa, para lo cual dividiremos cada frecuencia absoluta entre el número total de encuestados y lo multiplicamos por 100 para obtener el resultado en porcentajes, en el sabor chocolate tenemos lo siguiente:

$$\text{Frecuencia Relativa} = \frac{10}{60} \times 100 = 16.67\%$$

Una vez hecho lo mismo para todos los sabores tenemos:

Sabor	frecuencia absoluta	frecuencia relativa %	Grados
Chocolate	10	16.67	
Vainilla	16	26.67	
Fresa	13	21.67	
Queso de bola	21	35.00	
Total	60	100	

Siguiente paso es calcular cuantos grados de la gráfica circular corresponden a cada sabor que eligieron los encuestados, para ello dividiremos los 360 grados de un círculo entre el número de encuestados así tendremos lo siguiente:

Cobanota
La suma de todas las frecuencias relativas porcentuales debe ser igual a cero.

$$\text{Grados por persona encuestada} = \frac{360 \text{ grados}}{\text{número total de encuestados}} = \frac{360}{60} = 6$$

Lo que significa que 6 grados corresponde por cada persona encuestada, ahora calculamos los grados correspondientes por cada sabor elegida en la encuesta.

Para el sabor chocolate 10 personas eligieron ese sabor, lo que significa que tengo que multiplicar 6 grados por 10 personas lo que nos da lo siguiente:

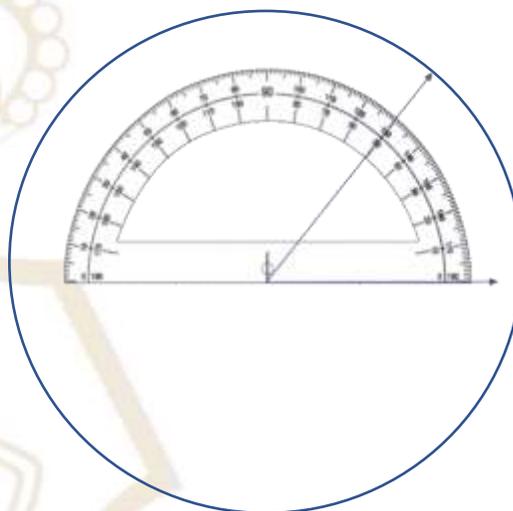
$$\text{Grados del sabor chocolate} = 6 \times 10 = 60 \text{ grados}$$

Hacemos lo mismo con los demás valores de la variable (sabores) y tenemos lo siguiente:

Sabor	frecuencia absoluta	frecuencia relativa %	Grados
Chocolate	10	16.67	60
Vainilla	16	26.67	96
Fresa	13	21.67	78
Queso de bola	21	35.00	126
Total	60	100	360

Una vez concluida la tabla procedemos a realizar la gráfica, para lo cual iniciamos dibujando un círculo.

Cobanota
 La suma de todos los grados nos debe dar 360 grados, en caso de no cumplir con esta suma algo hicimos mal por lo que deberemos revisar de nuevo los cálculos.



Con la ayuda de un transportador empezamos a trazar los grados correspondientes a cada sabor para darle forma a cada una de las rebanadas.

Al finalizar puedes darle color a cada sector, coloreando de diferente color, también deberás colocar un título al gráfico y las etiquetas que simbolizan cada sector con su respectivo color, así como el valor que corresponda, aquí lo más conveniente es etiquetar cada sector con el porcentaje que representa con respecto a todo el gráfico, este valor es la frecuencia relativa porcentual; para nuestro ejemplo el gráfico queda de la siguiente manera:



Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.

¿Quieres aprender a construir una gráfica circular?

Da clic al COBACHITO simultáneamente con la tecla CTRL



<https://youtu.be/RBgtRte7r5w>
Da clic a COBACHITO simultáneamente
con la tecla CTRL

GRÁFICA OJIVA

Para la construcción de una ojiva primero debemos adicionar una columna en la tabla de frecuencia que indiquen el número de observaciones cuyo valor sea menor o igual que el límite superior de cada clase anterior, lo que se conoce como frecuencia acumulada. Así, por ejemplo, siguiendo la tabla que a continuación se muestra, para la tercera clase, el número de observaciones menores a 189 son $17 + 13 + 2 = 32$, que son las frecuencias de la tercera, segunda y primera clase respectivamente. Este valor es la frecuencia acumulada hasta la tercera clase.

Tiempo invertido en atender clientes	Marca de clase	No. De clientes	Porción de clientes	Frecuencia absoluta acumulada	Frecuencia relativa acumulada
[141-157)	149	2	0.04	2	0.04
[157-173)	165	13	0.26	15	0.30
[173-189)	181	17	0.34	32	0.64
[189-205)	197	14	0.28	46	0.92
[205-221)	213	3	0.06	49	0.98
[221-237}	229	1	0.02	50	1.00
Total	50				

Si se grafican ahora sobre el eje de las X los límites superiores de clase y sobre el eje de las Y las frecuencias acumuladas absolutas o relativas obtenemos la gráfica conocida como Ojiva o Polígonos de frecuencia acumulada. Para este ejemplo la ojiva queda de la siguiente manera.



Gráfico elaborado por el Mtro. José Eduardo Mendoza Pérez, Plantel No. 2, mayo 2021.

Algunas de las afirmaciones que podemos hacer al observar la gráfica, son las siguientes:

- 46 clientes tardaron menos de 205 segundos; es decir el 92% de los clientes tardaron menos de 205 segundos.
- El proceso es muy tardado sólo en el 2% de los casos; tardaron más de 221 segundos.
- El proceso es muy rápido sólo en el 4% de los casos; tardaron menos de 157 segundos.
- El 64% de los clientes tardaron menos de 189 segundos.

2. Datos Agrupados con variable cuantitativa

Con el objetivo de saber cuánto es el **gasto aproximado de material didáctico semanal** de los estudiantes de 5to. semestre, se hizo una encuesta a los 80 alumnos de la serie de físico matemático. Considera los siguientes datos no agrupados y realiza la tabla de distribución de frecuencias. Obtén las representaciones correspondientes a la tabla.

Gasto semanal en material didáctico													
68.00	68.00	120.00	70.00	70.00	135.00	75.00	145.00	75.00	147.00	80.00	115.00	115.00	80.00
58.00	120.00	58.00	130.00	130.00	60.00	135.00	135.00	62.00	148.00	65.00	150.00	115.00	115.00
105.00	100.00	105.00	105.00	95.00	96.00	100.00	100.00	100.00	100.00	105.00	100.00	105.00	
81.00	82.00	83.00	85.00	86.00	81.00	85.00	90.00	90.00	92.00	92.00	93.00	150.00	
56.00	95.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	
108.00	107.00	107.00	115.00	115.00	115.00	115.00	115.00	115.00	115.00	115.00	158.00	115.00	

Intervalo		Marca de clase (mc)	Frecuencia (f)	Frecuencia relativa	Frecuencia acumulada (F)
Límite Inferior	Límite Superior				
			$\Sigma=$	$\Sigma=$	

Análisis de la Tabla de Frecuencia

3. Durante el mes de mayo, el servicio metrológico de nuestro país anunció que en la capital tabasqueña se registrarían temperaturas máximas récord en este año, por lo que el Colegio de Bachilleres de Tabasco, llevó un registro de las temperaturas a una misma hora durante todo este mes, los resultados se muestran a continuación:

Registro de temperaturas

41.5 42.8 43.3 43.5 45.1 46.2 43.7 45.2 44.5 46.6 47.6 48.7 46.1 48.5 49.8 47.5
48.3 50.8 48.3 47.4 49.1 47.3 44.6 43.7 46.7 48.9 51.1 51.0 46.8 49.1 48.2

a. Construye los intervalos correspondientes según la regla de Sturges y organiza los datos en una tabla de frecuencia agrupados por intervalos.

Intervalo		Marca de clase (mc)	Frecuencia (f)	Frecuencia relativa	Frecuencia acumulada (F)
Límite Inferior	Limite Superior				
Σ=			Σ=		

Análisis de la Tabla de Frecuencia

4. La cafetería del Plantel desea estandarizar la cantidad de paletas que deberá elaborar para su venta diaria, por lo que hizo un registro de las unidades vendidas durante los últimos 20 días, los resultados se muestran a continuación:

Registro de paletas vendidas por días									
25	28	26	25	29	30	31	26	28	25
27	27	26	28	29	27	28	31	26	28.

a. Organiza los datos en una tabla de frecuencia simple

Intervalo					
Límite Inferior	Limite Superior	Marca de clase (mc)	Frecuencia (f)	Frecuencia relativa	Frecuencia acumulada (F)
			$\Sigma=$	$\Sigma=$	

Análisis de la Tabla de Frecuencia

PYEI-B2-LC02 Lista de Cotejo para evaluar Tarea 02 Problemario

Asignatura:	Probabilidad y Estadística I	Bloque: II	Descripción básica de un conjunto de datos	Fecha:		
Nombres				Grupo:		
				Turno		
Ejercicios: Frecuencias para datos agrupados y no agrupados						
Aprendizajes Esperados			Contenidos Específico			
<p>Emplea las tablas de distribuciones de frecuencias para describir de manera crítica y reflexiva los resultados de investigaciones contextualizadas.</p>			<ul style="list-style-type: none"> ▪ Representación tabular mediante distribución de frecuencias. ▪ Determinación de clases para datos agrupados. 			
CRITERIOS				%	CUMPLE	Puntaje
					SI	NO
7. Entrega su producto terminado en el tiempo establecido por el facilitador				10%		
8. Identifica los datos de acuerdo a la población objetivo.				20%		
9. Organiza adecuadamente los datos en la tabla de distribución de frecuencias.				20%		
10. Aplica la regla de Sturges donde se solicita				20%		
11. Analiza adecuadamente la información obtenida en las distribuciones de frecuencias				20%		
12. Se relaciona colaborativamente mostrando disposición al trabajo metódico y organizado				10%		
Calificación						
Logros obtenidos			Aspectos a mejorar			
Nombre y Firma del Coevaluador			Firma del Facilitador			



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

Competencias a desarrollar		TAREA No. 3	EJERCICIOS: Frecuencias para datos agrupados y no agrupados
CG. 3.2 CG. 4.1 CG. 5.2 CG. 8.1	CDBM. 2 CDBM. 4 CDBM. 5 CDBM. 7 CDBM. 8		INSTRUCCIONES: En equipos de 6 según se integraron para la solución de la SD1, aplica el cuestionario siguiente con el propósito de recolectar datos y dar solución a la SD1 más adelante.

RECORDANDO PROPOSITO DE LA SD1

En equipos de 6 elementos **elaborar** las tablas y gráficas que representen las tallas de pantalones, faldas y playeras de los estudiantes de nuevo ingreso (**muestra del 20% de la población**), mediante la recolección de datos y el muestreo para generar la solución de la situación y presentarla ante el grupo para su socialización

En caso de ser mujer responde la respuesta 1 y 2, caso contrario la 3 y 4

1. ¿Cuál es tu talla de falda?

- a) 32 (CH)
- b) 34 (M)
- c) 36 (G)
- d) 38 (L)
- e) 40 (XL)
- f) 42 (XXL)

4. ¿Qué talla de camisa utilizas?

- a) 36
- b) 38
- c) 40
- d) 42
- e) 44

2. ¿Qué talla de blusa utilizas?

- a) 28 (CH)
- b) 30 (M)
- c) 32 (G)
- d) 34 (L)
- e) 36 (XL)
- f) 38 (XXL)

Nota. Se sugiere elaborar en un documento Forms

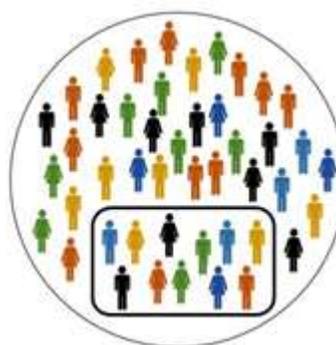
3. ¿Cuál es tu talla de pantalón?

- a) 28
- b) 30
- c) 32
- d) 34
- e) 36
- f) 38
- g) 40
- h) 42
- i) 44

PYEI-B1_B2-PP01		CUESTIONARIO TIPO PLANEA BLOQUE I Y II
CG 1.1 CG 4.1 CG 5.1 CG 5.2 CG 5.6 CG 8.2	CDBM 1 CDBM 2 CDBM 4 CDBM 5 CDBM 8	Indicaciones Lee cuidadosamente cada reactivo e individualmente elige la opción de respuesta correcta

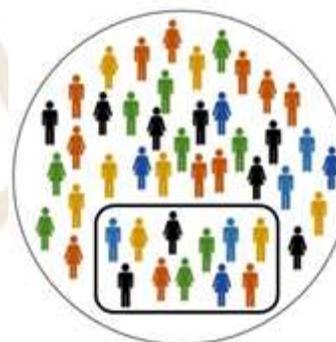
1. Tomando en cuenta la siguiente imagen los elementos dentro la circunferencia representan...

- a) Población
- b) Muestra
- c) Dato
- d) Muestreo



2. Tomando en cuenta la siguiente imagen los elementos dentro del rectángulo representan...

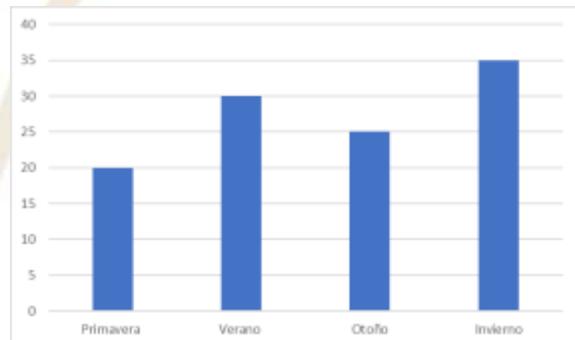
- a) Población
- b) Muestra
- c) Dato
- d) Muestreo



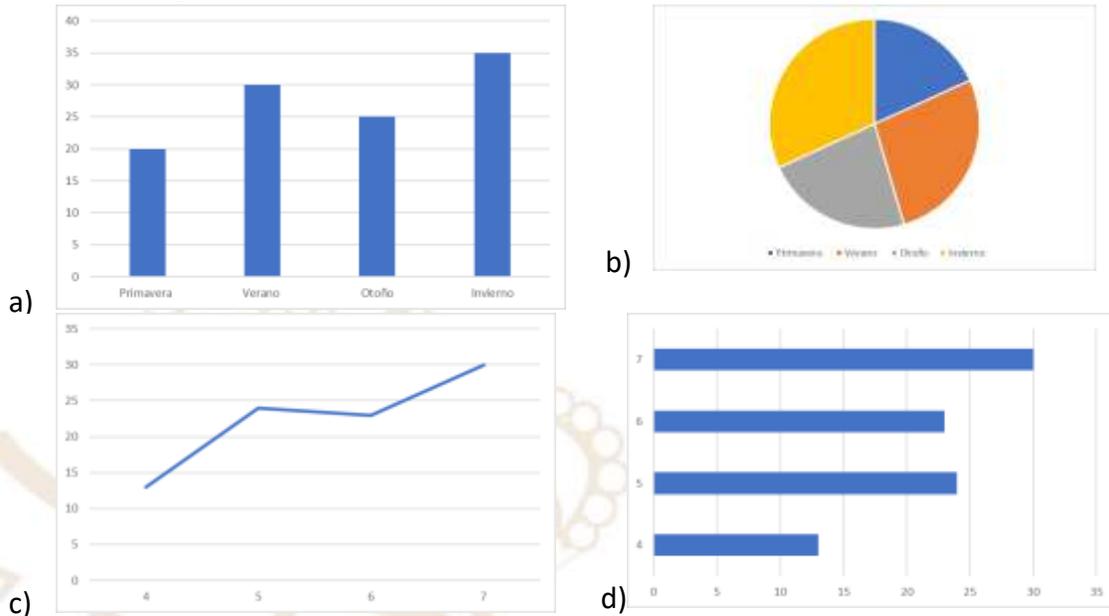
3. Toda porción de elementos tomados de un universo.

- a) Población
- b) Muestra
- c) Dato
- d) Muestreo

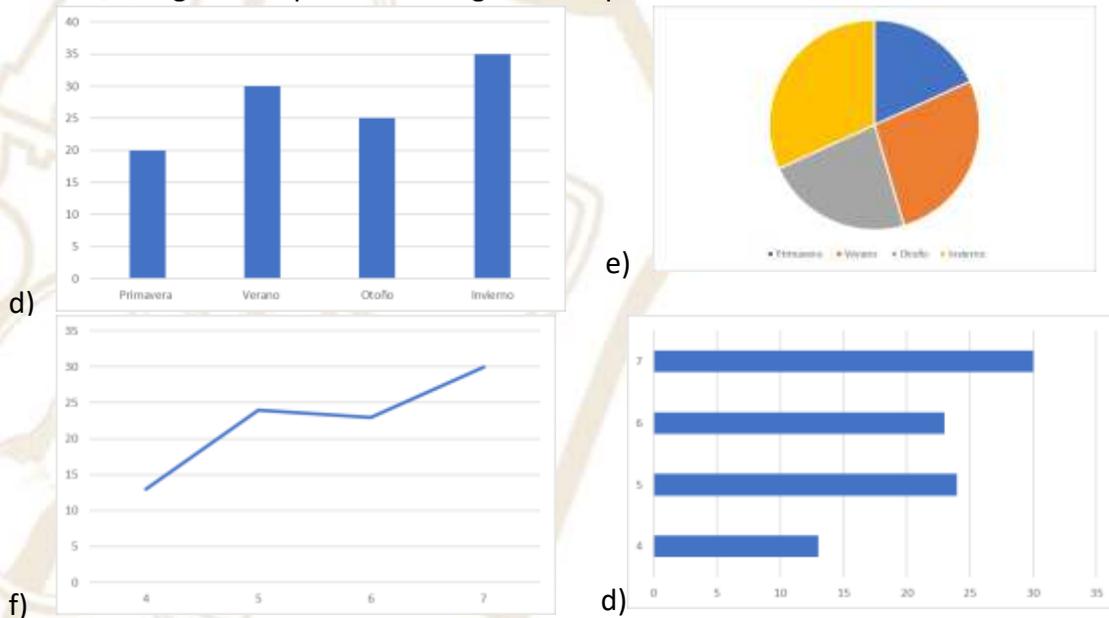
4. "Color de ojos" es una variable de tipo:
- Cualitativa nominal
 - Cualitativa ordinal
 - Cuantitativa Discreta
 - Cuantitativa continua
5. "Gusto por las matemáticas" es una variable de tipo:
- Cualitativa nominal
 - Cualitativa ordinal
 - Cuantitativa Discreta
 - Cuantitativa continua
6. "Edad" es una variable de tipo:
- Cualitativa nominal
 - Cualitativa ordinal
 - Cuantitativa Discreta
 - Cuantitativa continua
7. "Estatura" es una variable de tipo:
- Cualitativa nominal
 - Cualitativa ordinal
 - Cuantitativa Discreta
 - Cuantitativa continua
8. Tomando en cuenta la siguiente grafica que representa la altura de las barras.
- Frecuencia
 - población
 - Muestra
 - Variable



9. ¿Qué grafica representa un polígono de frecuencias?



10. ¿Qué grafica representa una gráfica de pastel?



PYEI-B1_B2-MA01 Mapa de aprendizaje para evaluar los Aprendizajes Esperados

Asignatura:	Probabilidad y Estadística I	Bloque: I y II	Elementos Estadísticos y Descripción Gráfica De Un Conjunto De Datos	Fecha:	
Nombre				Grupo:	
				Turno:	

Situación Didáctica 1: "Uniforma e informa"

Conocimientos	Habilidades	Actitudes
<p>Elementos estadísticos</p> <ul style="list-style-type: none"> • Población • Dato • Muestra • Variable y sus tipos <p>Técnicas de recolección de datos</p> <ul style="list-style-type: none"> • Encuesta • Entrevista • Observación • Experimentación • Documental <p>Técnicas de muestreo</p> <ul style="list-style-type: none"> • Probabilístico • No probabilístico 	<ul style="list-style-type: none"> • Distingue los elementos estadísticos, el tipo de población y sus características • Ordena datos por medio de diferentes técnicas de recolección • Diferencia las diversas formas para elegir una muestra 	<ul style="list-style-type: none"> • Se relaciona con sus semejantes de forma colaborativa mostrando disposición al trabajo metódico organizado. • Escucha y participa activamente • Se informa a través de diferentes fuentes antes de tomar decisiones • Reflexiona sobre diferentes posturas de conducirse en el contexto
<p>Representación tabular mediante distribución de frecuencias</p> <ul style="list-style-type: none"> • Frecuencia absoluta • Frecuencia acumulada • Frecuencia relativa <p>-Determinación de clases para datos agrupados</p> <p>-Criterios, representación y análisis de graficas</p> <ul style="list-style-type: none"> • Histograma • Polígono de frecuencias • Circular • Ojiva 	<ul style="list-style-type: none"> • Identifica las características de la población objetivo mediante la distribución de frecuencias • Reconoce los pasos a seguir en la elaboración de tablas de distribución de frecuencias • Ilustra gráficamente las variables obtenidas 	<ul style="list-style-type: none"> • Expresa ideas y conceptos favoreciendo su creatividad. • Toma decisiones de manera consciente e informada asumiendo las consecuencias • Es tolerante con posturas diferentes a las suyas • Externa un pensamiento crítico y reflexivo de manera solidaria



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

- Analiza la información obtenida en las distribuciones de frecuencias.

1 = Necesito ayuda

2 = Puedo hacerlo solo

3 = Puedo ayudar a otros

CRITERIOS

NIVEL

1 2 3

¿Qué debo hacer para mejorar?

Aplico crítica y reflexivamente los elementos estadísticos, así como la información recolectada a través de las diferentes técnicas, su tipo de variabilidad y su comportamiento en diferentes contextos

Empleo las tablas de distribuciones de frecuencia para describir de manera crítica y reflexiva los resultados de investigaciones contextualizadas.

Utilizo las gráficas como un medio creativo para comparar valores y facilitar la toma responsable de decisiones en problemas presentes en cualquier contexto

Nombre y Firma del Coevaluador

Firma del Facilitador

Referencias BLOQUE I Y II

Díaz, F. (s. f.). *Elementos de muestreo*. <http://cienciassociales.webcindario.com/PDF/TecMuestreo.pdf>. Recuperado 18 de mayo de 2021, de <http://cienciassociales.webcindario.com/PDF/TecMuestreo.pdf>

Gonzalez, W. (13 de 05 de 2009). <http://recodatos.blogspot.com/2009/05/tecnicas-de-recoleccion-de-datos.html>. Obtenido de <http://recodatos.blogspot.com/2009/05/tecnicas-de-recoleccion-de-datos.html>: <http://recodatos.blogspot.com/2009/05/tecnicas-de-recoleccion-de-datos.html>

Daniel Carreón. (2021, 23 febrero). *TABLA DE FRECUENCIAS AGRUPADA EN INTERVALOS Super facil para principiantes*. YouTube. <https://www.youtube.com/watch?v=VNMck8wco98>

Guerrero, L. S. J. (2016). *Probabilidad y estadística: Para bachilleratos tecnológicos*. Grupo Editorial Patria.

Cómo hacer una tabla de frecuencias paso a paso. Ejemplos resueltos. (2021, 26 enero). Clases de Matemáticas Online. <https://ekuatío.com/como-hacer-una-tabla-de-frecuencias-paso-a-paso-ejemplos-resueltos/>

Matemóvil. (2018, 20 agosto). *Tabla de Frecuencias para Datos no Agrupados - Ejercicios Resueltos*. YouTube. <https://www.youtube.com/watch?v=iPEt789ewVM>

Ortiz Cruz, I., Gómez Rodríguez, E., & Martínez Sánchez, L. (2016, agosto). *Componente Formaci3n Propedéutica - Google Drive*. COBACHBC. https://drive.google.com/drive/folders/1FIJ_44yu5Jj9iyBN5MVqlgqgAbeqWiFS



René Jiménez, M., Jiménez Gallegos, J. A., & Estrada Coronado, R. M. (2020). *PROBABILIDAD Y ESTADÍSTICA* (Primera Edición ed.). Pearson Education.

Carreón, D. (2017, 2 agosto). *COMO HACER UNA GRÁFICA CIRCULAR Super fácil*. YouTube. <https://www.youtube.com/watch?v=RBgtRte7r5w&feature=youtu.be>

E. (2015, 31 agosto). *¿Qué es un histograma?* YouTube. <https://www.youtube.com/watch?v=XFjDGydsGvU&feature=youtu.be>

fisterra.com. (2021). *Metodología investigación: Representación gráfica en el Análisis de Datos*. Metodología de la Investigación. <https://www.fisterra.com/formacion/metodologia-investigacion/representacion-grafica-analisis-datos/index.asp>

M. (2018, 20 agosto). *Tabla de Frecuencias para Datos Agrupados en Intervalos - Ejercicios Resueltos*. YouTube. <https://www.youtube.com/watch?v=6ygaz0vECzY&feature=youtu.be>

Probabilidad y Estadística 1 (2.ª ed.). (2016). [Libro electrónico]. https://qrcgcustomers.s3-eu-west-1.amazonaws.com/account13675116/14845856_2.pdf?0.648130073595317

Redondeando Números Enteros. (2021). montereyinstitute. https://www.montereyinstitute.org/courses/DevelopmentalMath/TEXTGROUP-1-8_RESOURCE/U01_L1_T2_text_final_es.html#:~:text=Normalmente%2C%20



se%20redondea%20al%20n%C3%BAmero,de%20190%2C%20redondear%20A
Das%20a%20190

Rodó, P. (2021, 12 enero). *Gráfico circular*. Economipedia.

<https://economipedia.com/definiciones/grafico-circular.html>

S. (2020, 8 septiembre). *Polígono de Frecuencias - Qué es, para qué sirve y cómo hacer uno*.

Concepto. <https://concepto.de/poligono-de-frecuencias/#ixzz6vEeqqm3V>

Torres, R. (2020, 1 diciembre). *Histograma -Definición, pasos y formas típicas*.

Soloindustriales. <https://soloindustriales.com/histograma/>

(Flaticon, 2021)

(Freepik, 2020)



BLOQUE III: Medidas estadísticas

Propósito del bloque III



Bloque IV

Demuestra generalizaciones de comportamientos, así como sesgos en diferentes poblaciones objetivo de cualquier ámbito presente en su entorno, por medio del uso creativo de las medidas de tendencia central y dispersión.

Aprendizajes Esperados

- Examina de manera crítica información resumida, por medio de las medidas de tendencia central obtenidas en diferentes contextos.
- Emplea comparaciones entre la dispersión de datos que le permitan estimar su comportamiento producto de una investigación y los datos que arrojan las medidas estadísticas para la toma consciente de decisiones.
- Usa límites estadísticos para el análisis crítico y reflexivo de situaciones problema.



Competencias	
Genéricas	Disciplinares
<p>5. Desarrolla innovaciones y propone soluciones a problemas a partir de métodos establecidos</p> <p>5.1 Sigue instrucciones y procedimientos de manera reflexiva, comprendiendo como cada uno de sus pasos contribuye al alcance de un objetivo.</p> <p>5.5 sintetiza evidencias obtenidas mediante la experimentación para producir conclusiones y generar nuevas preguntas</p> <p>8. participa y colabora de manera efectiva en equipos diversos</p> <p>8.3 asume una actitud constructiva, congruente con los conocimientos y habilidades con los que cuenta dentro de diferentes equipos de trabajo</p>	<ol style="list-style-type: none">1. Construye e interpreta modelos matemáticos mediante la aplicación de procedimientos aritméticos, algebraicos geométricos y variacionales para la comprensión y análisis de situaciones reales hipotéticas o formales2. Formula y resuelve problemas matemáticos aplicando diferentes enfoques3. Explica e interpreta los resultados obtenidos mediante procedimientos matemáticos y los contrasta con modelos establecidos o situaciones reales4. Argumenta la solución obtenida de un problema, con métodos gráficos analíticos o variacionales, mediante el lenguaje verbal, matemático y el uso de las tecnologías de la información y la comunicación.

Situación Didáctica 2

Título: "EnTALLA2"

La empresa fabricante de uniformes escolares que establecerá convenio con el COBATAB, desea expandir su colaboración con otras instituciones educativas del estado. Para aprovechar al máximo su materia prima en la elaboración de playeras, requiere saber las medidas del largo de los pliegues (cm promedio que definen la talla) que necesitan cortar. Tomando como base la siguiente tabla:

ROPA							
	PECHO		CINTURA		CADERA		
	PLG	CM	PLG	CM	PLG	CM	
 HOMBRE	CH	36-38	92-96	29-31	74-79	36-38	92-96
	M	38-40	96-102	31-33	79-84	38-40	96-102
	G	40-42	102-106	33-35	84-89	40-42	102-106
	XG	43-45	109-114	36-38	92-96	42-44	106-112
 MUJER	XCH	30-32	78-83	23-25	58-63	33-35	85-89
	CH	32-35	83-89	25-27	63-68	35-37	90-95
	M	35-37	89-94	27-28	68-71	37-39	95-101
	G	37-39	95-101	29-32	74-81	39-42	101-108
	XG	40-42	102-108	31-33	81-87	42-44	109-113

<https://i.pinimg.com/564x/63/2f/00/632f00b16dd6fae744fd12d0dbadcf4.jpg>

Conflicto cognitivo:

- A) ¿Cuál es la medida (cm) de los pliegues que más necesitan los fabricantes de dicha empresa?
- B) ¿Cuál es el sesgo en la distribución de los datos obtenidos?
- C) Interpreta tus resultados.

Propósito de la situación didáctica 2

En equipos de 6 elementos, apoyados en la tabla de la imagen que hace referencia a la talla de playeras y la medida en cm del pecho de una persona, **construir una nueva tabla de distribución de frecuencias, de las tallas de playeras de los estudiantes de nuevo ingreso, y elaborar un reporte donde determine e interprete las medidas de tendencia central que resuelven las preferencias favorables para la empresa en el corte de los pliegues, así como su variabilidad y presentarla ante el grupo para su socialización.**



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

Instrumento de Evaluación Situación Didáctica 2

PYE1-B1-B2-LC03 Lista de cotejo para evaluar la Situación Didáctica 2

INSTRUMENTO DE EVALUACION						
LISTA DE COTEJO						
SITUACION DIDÁCTICA "EnTALLA2"						
DATOS GENERALES						
Nombre(s) del alumno(s)			Matricula(s)			
Producto: Documento electrónico y/o impreso			Fecha			
Materia: Hoja de Cálculo Aplicado			Periodo: 20'B			
Nombre del docente			Firma del docente			
CRITERIO	INDICADORES	VALOR OBTENIDO		PONDERACIÓN	CALIF	OBSERVACIONES Y/O SUGERENCIAS DE MEJORA
		SI	NO			
1	Examina y organiza de manera crítica información resumida, por medio del uso de tablas de frecuencia.			15		
2	Utiliza y calcula acertadamente las medidas de tendencia central obtenidas en la solución de la situación planteada.			25		
3	Emplea comparaciones entre la dispersión de datos que le permitan estimar su comportamiento producto de una investigación.			10		
4	Expresa una toma consciente de decisión sobre el comportamiento que arrojan los datos según las medidas estadísticas y los límites estadísticos.			20		
5	Trabaja colaborativamente en el proceso de solución de la situación didáctica.			10		
6	Entrega en tiempo y forma el reporte de la situación didáctica.			10		
7	Utiliza las TIC's como apoyo en la elaboración de las gráficas o tablas utilizadas.			10		
				CALIFICACION		

Actividad Construye T

9.5 ¿Cómo saber si la información es confiable?

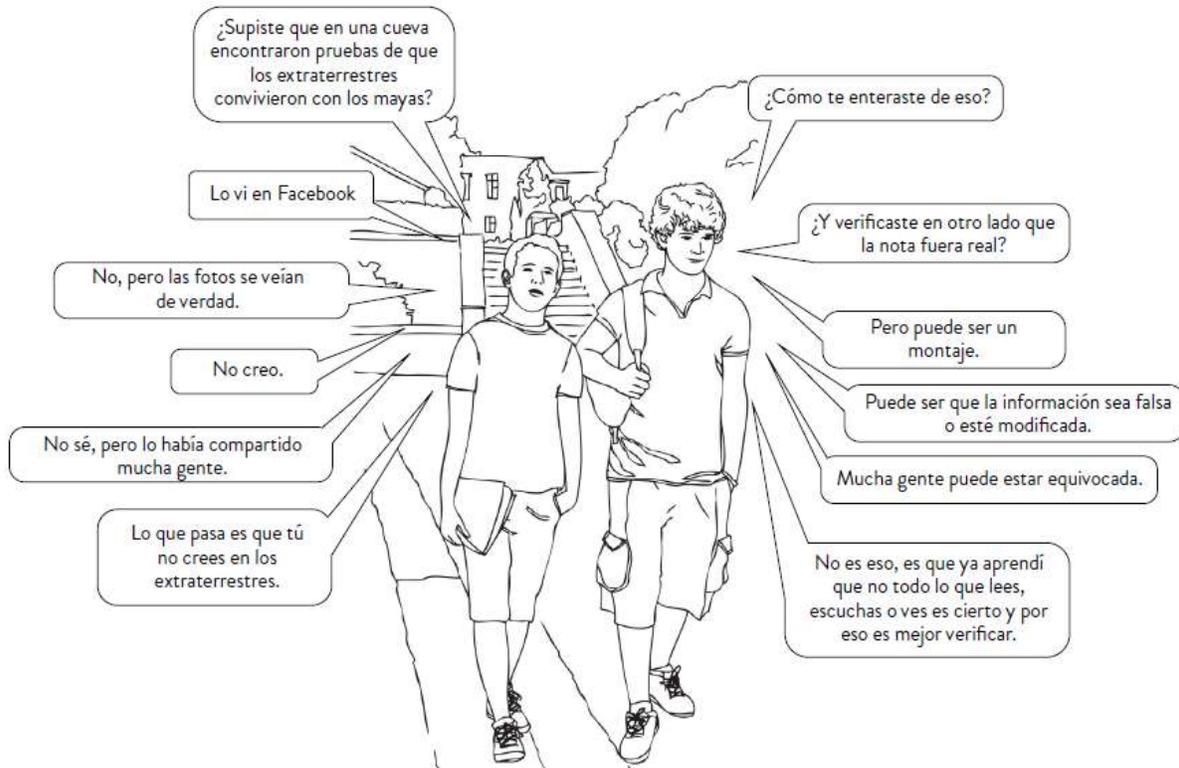
"Creer o verificar, la alternativa es ineludible".
Kuan - Tseu.

La información que recibimos de las experiencias, el entorno social próximo, el contexto, las fuentes documentales y los medios de comunicación suele ser de utilidad para orientar las decisiones, sin embargo, siempre es conveniente verificar que sea confiable, real y significativa. Para ello, existen estrategias que evalúan los datos recibidos y permiten la activación de un proceso de selección para aprovechar lo que es confiable y descartar lo que no cumple con este criterio.

El reto es contrastar información relevante de diferentes fuentes, experiencias previas y redes de apoyo para facilitar la toma de decisiones ante diversas alternativas de acción.

Actividad 1

a. En parejas, observen la imagen:



- b. Con base en el diálogo, respondan las siguientes preguntas:
- ¿Qué posibilidades tienen cada una de las personas que participan en el diálogo de tomar decisiones responsables?
 - ¿Les ha pasado que han confiado en la información que reciben sin verificarla?
 - ¿La información falsa, los rumores o las opiniones centradas en prejuicios o creencias sin sustento pueden afectar las decisiones de alguien? ¿Por qué?
- c. Compartan con el grupo sus conclusiones.

Actividad 2

- a. Individualmente, elabora una ruta que incluya lo que haces para asegurarte de que la información que recibes es confiable. Responde aquí o copia el esquema en tu cuaderno:



- b. Comparte tu ruta con un compañero.

Escribe en un minuto
qué te llevas de la lección



Reafirmo y ordeno

Contrastar la información verificando en distintas fuentes, personas, instituciones o recursos, permite que contemos con más certezas para el manejo y aprovechamiento de dicha información para la toma responsable de decisiones.

Para tu vida diaria

Cada vez que recibas una noticia, busques información en internet o escuches algo sobre alguien, verifícalo antes de asumir ese dato como verdadero. Busca una noticia que se repita en tres o cuatro periódicos y contrasta la información, para identificar si los datos coinciden.

¿Quieres saber más?

Revisa el video: "¿Es verdad lo que encontramos en la web?", en el que se exponen estrategias para evaluar la confiabilidad de la información que se encuentra en internet. Búscalo en tu navegador o entra en la siguiente dirección: <https://youtu.be/B3Z6wI2Ds7A>

Concepto clave

Información confiable: Aquella información que está respaldada por una institución reconocida o tiene como soporte teorías o métodos propios de la investigación científica.

Evaluación Diagnóstica

¿Qué tanto sé? (Apertura)



PYEI-B3-ED02

Evaluación diagnóstica "Medidas Estadísticas"

BLOQUE III: Medidas estadísticas

NOMBRE

GRUPO

FECHA

Instrucciones: Lee cuidadosamente cada una de las siguientes cuestiones y elige la opción de respuesta correcta.

1.- Son valores que nos indican alrededor de que valor se agrupa el mayor número de observaciones de la variable, por lo que son representativas de todos los datos.

- a) Medidas de Tendencia Central
- b) Medidas de Dispersión
- c) Medidas de Posición
- d) Medidas de Tanteo

2.- Para ampliar nuestro conocimiento del patrón de los datos, debemos calcular también la variabilidad o extensión. Esto lo podemos hacer a través de:

- a) Medidas de Tendencia Central
- b) Medidas de Dispersión
- c) Medidas Automáticas
- d) Medidas de Tanteo

3.- Corresponde al promedio de todas las observaciones.

- a) Mediana
- b) Media
- c) Moda
- d) Desviación estándar (s)

4.- Es el valor que ocupa la posición central en un conjunto ordenado de observaciones.

- a) Mediana
- b) Media
- c) Moda
- d) Desviación estándar (s)

5.- Es el valor que más se repite dentro del conjunto de datos.

- a) Mediana
- b) Media
- c) Moda
- d) Desviación estándar (s)

6.- Es la raíz cuadrada del promedio de los cuadrados de las diferencias de cada dato con respecto de la media.

- a) Mediana
- b) Media
- c) Moda
- d) Desviación estándar (s)



TABASCO

"Educación que genera cambio"



COBATAB

COLEGIO DE BACHILLERES DE TABASCO

7.- Es el promedio de los valores absolutos de las diferencias de cada dato con respecto de la media.

- a) Rango
- b) Desviación Media
- c) Varianza (S^2)
- d) Desviación estándar (s)

8.- Es la diferencia entre el valor mayor y el menor de un conjunto de datos.

- a) Rango
- b) Desviación Media
- c) Varianza (S^2)
- d) Desviación estándar (s)

9.- Es el promedio de los cuadrados de las diferencias de cada dato con respecto de la media.

- a) Rango
- b) Desviación Media
- c) Varianza (S^2)
- d) Desviación estándar (s)

10.- Además de la mediana, existen otros datos estadígrafos que te permiten dividir tus datos ordenados en partes. Estos pueden ser:

- a) Cuartiles (Q), Deciles (D) y Percentiles (P)
- b) Cuartiles (Q), Rango (R) y Percentiles (P)
- c) Cuartetos (Q), Mediana (M), Varianza (V)
- d) Cuartetos (Q), Deciles (D) y Varianza (V)

11.- Son aquellas entidades numéricas utilizadas para señalar la posición que ocupa un dato determinado, en relación con el resto de los datos numéricos, permitiendo así conocer otros puntos propios de la distribución de datos, que no son inherentes a los valores centrales.

- a) Desviación Estándar estadística
- b) Límites estadísticos
- c) Límites y sublímites
- d) Límite de desviación

12.- Las calificaciones de un estudiante en una prueba fueron las siguientes: Química 85, español 91, Matemáticas 78, Historia 82, inglés 79, Ética 95. ¿Cuál es el promedio correspondiente?

- a) 82 y 85
- b) 82
- c) 85
- d) 510

13.- Cuál es la mediana de los siguientes datos: 10, 2, 5, 4, 7 y 3

- a) 4
- b) 5
- c) 10
- d) 2

14.- Cuál es la moda de los siguientes datos: 3, 4, 4, 6, 7, 7, 9, 11

- a) 4
- b) 7
- c) 4; 7
- d) No hay moda

15.- Cuál es la desviación media de la distribución: 9, 3, 8, 8, 9, 8, 9, 18

- a) 9
- b) 8
- c) 8.5
- d) 2.5



Conceptos Básicos

Medidas de tendencia central

Las medidas de tendencia central son valores estadísticos que nos permiten interpretar información a partir de un conjunto de valores.

Para un conjunto de observaciones en una distribución de valores, las medidas de tendencia central se encuentran hacia el centro de la distribución y entre ellas se considera a la Media, Mediana y Moda. Dichas medidas pueden calcularse a cualquier conjunto de variables, generalmente cuantitativas, ya sea como un conjunto de datos agrupados o no agrupados.

¿Datos no agrupados?

Es una muestra o conjunto de datos que no ha sido organizado y se presentan de forma individual o en listas.

Ejemplos:

1. Las calificaciones de 10 alumnos en una clase de Probabilidad		5, 6, 8, 8, 9, 10, 5, 7, 8, 10
2. El peso de 17 personas en kilogramos		66, 65, 59, 82, 64, 55, 76, 64, 67, 71, 48, 52, 65, 69, 80, 58, 65

¿Datos agrupados?

Es una muestra o conjunto de datos que han sido organizados o clasificados en clases o intervalos.

Ejemplo:

Edad (x)	Marca de Clase (x_i)	Frecuencia Absoluta (f_i)	Frecuencia Absoluta Acumulada (F_i)
[10-19)	14.5	5	5
[19-28)	23.5	11	16
[28-37)	32.5	8	24
[37-46)	41.5	5	29
[46-55)	50.5	8	37
[55-64)	59.5	6	43
[64-73)	68.5	7	50
	Total	50	



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

Media

También es llamada **promedio** de un conjunto de "n" observaciones $x_1, x_2, x_3, \dots, x_n$. La media aritmética se representa con el símbolo \bar{x} .

La media aritmética se obtiene sumando todas las observaciones y dividiendo dicha suma entre número total de elementos del conjunto de datos.

La fórmula matemática que se utiliza para calcular el valor de la media se expresa de la siguiente forma:

Datos Agrupados

$$\bar{x} = \frac{\sum x_i * f_i}{n}$$

Donde:

Σ : Sumatoria

x_i : Marca de clase

f_i : Frecuencia absoluta

n : Número total de datos

Datos no Agrupados

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

Donde:

$x_1 + x_2 + x_3 + \dots + x_n$: es la Suma total de los datos de la muestra

n : Número total de datos

Mediana

La mediana de un conjunto de "n" observaciones $x_1, x_2, x_3, \dots, x_n$ es el valor que ocupa la posición media de los datos ordenados de menor a mayor y se representa con el símbolo \tilde{x} .

La fórmula matemática que se utiliza para calcular la posición de la mediana se expresa de la siguiente forma:

Datos Agrupados

$$\tilde{x} = L_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} * A_i$$

Donde:

n = Número total de datos

L_i = Limite inferior de la clase de la mediana

F_{i-1} = Frecuencia absoluta acumulada anterior a la clase de la mediana

f_i = Frecuencia absoluta

A_i = Amplitud del intervalo

Datos no Agrupados

Muestra impar:

$$\tilde{x} = \frac{n + 1}{2}$$

Muestra par:

$$\tilde{x} = \frac{n}{2}$$

Donde:

n = Número total de datos

Para calcular la mediana de un conjunto de datos no agrupados debemos tener en cuenta lo siguiente:

1. Si el conjunto o muestra es **par** entonces la mediana será el promedio de los valores numéricos de las dos observaciones que se encuentran justo a la mitad del conjunto ordenado.
2. Si el conjunto o muestra es **impar** entonces el valor de la mediana es correspondiente a la posición indicada por la aplicación de la fórmula.

Moda

La moda dentro de un conjunto de observaciones $x_1, x_2, x_3, \dots, x_n$, es el valor que se repite con mayor frecuencia y se representa con el símbolo \hat{x} .

La moda dentro de un conjunto de datos puede darse a través de tres situaciones:

1. Si la moda no existe se denomina **amodal**.
2. Si existen dos modas dentro del conjunto de n observaciones se denomina **bimodal**.
3. Si en el conjunto de datos existen tres modas o más se denomina **multimodal**.

La fórmula matemática que se utiliza para calcular la moda se expresa de la siguiente forma:

$$\hat{x} = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} * A_i$$

Donde:

L_i = Límite inferior de la clase con mayor frecuencia.

f_i = Frecuencia absoluta del intervalo con mayor frecuencia.

f_{i-1} = Frecuencia absoluta anterior al intervalo de mayor frecuencia.

f_{i+1} = Frecuencia absoluta posterior al intervalo de mayor frecuencia.

A_i = Amplitud del intervalo

$$\hat{x} = \text{Dato que más se repite.}$$

Sesgo

También conocida como medidas de asimetría, esta medida nos permite identificar si dentro de un conjunto de observaciones $x_1, x_2, x_3, \dots, x_n$, los datos se distribuyen de manera uniforme alrededor del punto central o de la media aritmética y su representación gráfica se basa en las medidas de tendencia central (Media, Mediana y Moda), donde la moda siempre será representada por la línea más alta, por ser el dato que más se repite, la mediana es representada por una línea un poco más baja y al centro por ser el punto central y finalmente en la última posición se encuentra la media.

La gráfica de asimetría puede ser negativa (a la izquierda) o positiva (a la derecha), de igual forma puede ser simétrica, es decir, que la cantidad de datos sea el mismo número de valores tanto a la derecha como a la izquierda del punto central (Media, mediana y moda tienen el mismo valor).

Ejemplos:

1. Se realiza una encuesta a 13 personas acerca del número de hermanos que tienen, los datos obtenidos se muestran a continuación:

3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 6, 6.

Media= 4.08 Mediana= 4 Moda= 3

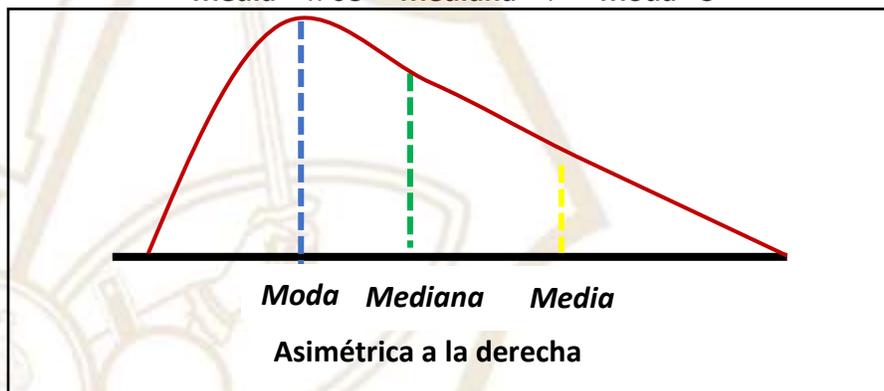


Gráfico elaborado por la docente Nancy Arias C., plantel 4, COBATAB, junio, 2021.

2. Se desea conocer las edades de los 14 niños que integran el equipo de futbol de la primaria Leonor Pontes de Jiménez, la información obtenida se presenta a continuación:

6, 7, 7, 8, 8, 8, 8, 9, 9, 9, 9, 9, 9, 9.



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

Media= 8.2 Mediana= 8.5 Moda= 9

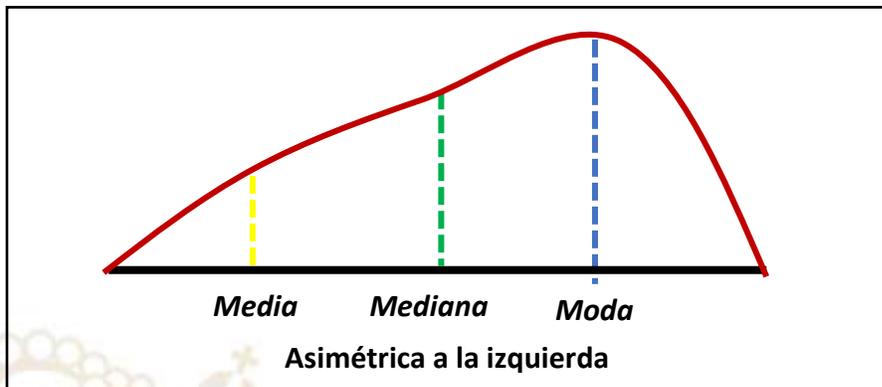


Gráfico elaborado por la docente Nancy Arias C., plantel 4, COBATAB, junio, 2021.

- Se han registrado los gastos mensuales de una escuela y al finalizar el análisis estadístico, se obtiene que el valor de la media es de \$200,000, la mediana es de \$200,000 y la moda es de \$200,000.

Como el valor de las medidas de tendencia es la misma para cada caso, nos encontramos con una distribución simétrica, como se muestra a continuación:

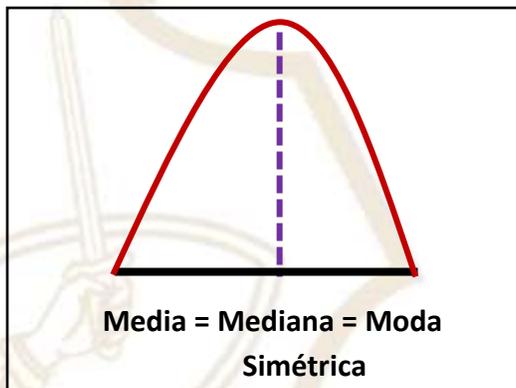


Gráfico elaborado por la docente Nancy Arias C., plantel 4, COBATAB, junio, 2021.

VIDEOS DE APOYO

YOUTUBE:

Distribución de Datos
Asimétrica



<https://www.youtube.com/watch?v=nb6NWlamFzY>

YOUTUBE:

Distribución de Datos Simétrica o
Campana de Gauss



<https://www.youtube.com/watch?v=WONJ6hztQc>

Medidas de Tendencia Central para Datos no Agrupados

Ejemplo 1

En una fábrica de ropa se ha tenido que modificar la cantidad de personal autorizado para laborar, esto debido a la pandemia Covid-19 que vivimos actualmente. De los 18 trabajadores que laboran en la fábrica únicamente deben laborar 9 personas por día. Como apoyo al personal, la empresa les permite laborar horas extras, a continuación, se muestran las horas extras que acumularon los 9 trabajadores en su última quincena de trabajo:

16, 19, 19, 24, 25, 26, 30, 30, 30.

Calcule la media, mediana y moda.

Media

Para obtener la media o el promedio, se suman cada uno de los datos de la muestra y se divide entre el número total de datos:

$$\bar{x} = \frac{16 + 19 + 19 + 24 + 25 + 26 + 30 + 30 + 30}{9}$$

$$\bar{x} = \frac{219}{9}$$

$$\bar{x} = 24.33 \text{ horas}$$

Mediana

Para obtener la mediana primero se deben ordenar los datos, sin embargo, como se puede observar el conjunto de la muestra ya se encuentra ordenado.

16, 19, 19, 24, **25**, 26, 30, 30, 30.

$$\tilde{x} = \frac{9 + 1}{2} = 5$$

El lugar de la mediana está en la posición 5, por lo tanto, el valor es:

$$\tilde{x} = 25 \text{ horas}$$

Moda

Recordemos que la moda dentro de un conjunto de datos no agrupados es el dato que más se repite, en este caso el valor de moda es 30.

16, 19, 19, 24, 25, 26, **30, 30, 30**.

$$\hat{x} = 30 \text{ horas}$$

Medidas de sesgo o asimetría

En este caso, el comportamiento de las medidas de tendencia central, tiene un sesgo asimétrico a la izquierda, como se observa en la imagen:

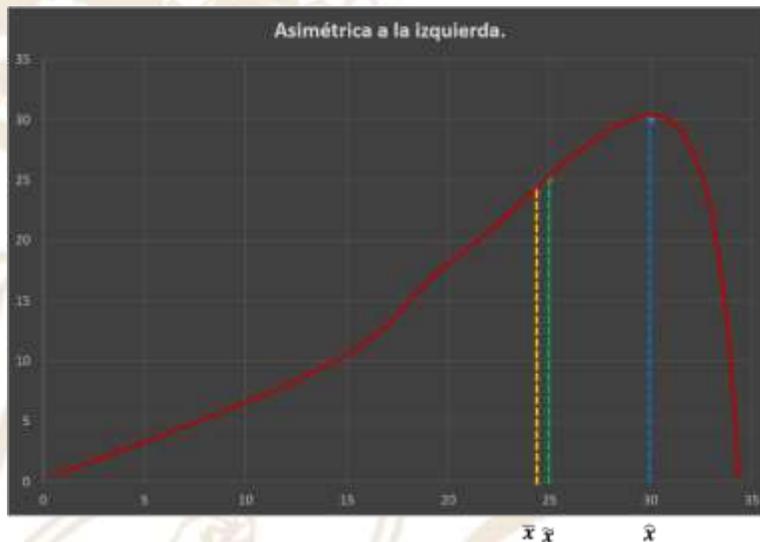


Gráfico elaborado por la docente Nancy Arias C., plantel 4, COBATAB, junio, 2021.

Ejemplo 2:

Diez jóvenes compitieron en la carrera de 200 metros planos. Los siguientes datos representan los tiempos, en segundos, que necesitaron para realizar el recorrido:

24, 32, 30, 26, 32, 24, 32, 22, 28, 30

Calcule la media, mediana y moda.

Media

Para esta colección de datos $n = 10$, y se calcula la media mediante la fórmula establecida.

$$\bar{x} = \frac{24 + 32 + 30 + 26 + 32 + 24 + 32 + 22 + 28 + 30}{10}$$

$$\bar{x} = \frac{280}{10}$$

$$\bar{x} = 28 \text{ segundos}$$

Mediana

Para obtener la mediana es importante recordar que primero se deben ordenar los datos de menor a mayor o viceversa.

Conjunto ordenado: 22, 24, 24, 26, **28, 30**, 30, 32, 32, 32.

$$\text{posición para } \tilde{x} = \frac{10}{2} = 5$$

El lugar de la mediana está en la posición 5, sin embargo, al ser una muestra con un conjunto de datos pares se toma el promedio de los valores que están en la posición 5 y la posición 6, en este caso la mediana es 29:

$$\tilde{x} = \frac{28 + 30}{2} = 29$$

$$\tilde{x} = 29$$

Moda

En este caso el valor de moda es:

22, 24, 24, 26, 28, 30, 30, **32, 32, 32**.

$$\hat{x} = 32$$

Medidas de sesgo o asimetría

Para este ejemplo el sesgo también resultó ser asimétrico a la izquierda.

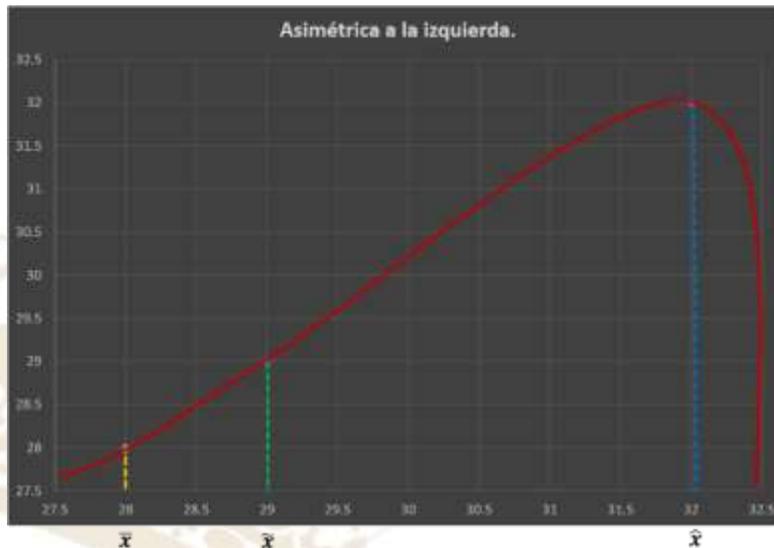


Gráfico elaborado por la docente Nancy Arias C., plantel 4, COBATAB, junio, 2021.

Medidas de Tendencia Central para Datos Agrupados

Ejemplo 1

En la siguiente tabla se reportan las edades de 20 estudiantes que asisten a un curso especial de Matemáticas.

Edad (x)	Marca de Clase (x_i)	Frecuencia Absoluta (f_i)	Frecuencia Absoluta Acumulada (F_i)
13 – 15	14	4	4
15 – 17	16	9	13
17 – 19	18	3	16
19 – 21	20	3	19
21 – 23	22	1	20
Total		20	

A partir de los datos obtenidos, hallar la media, mediana y moda de la muestra.

Media

Como sabemos, para hallar la media, es necesario promediar los datos de la muestra. En un conjunto de datos agrupados es necesario previamente calcular el promedio de la marca de clase, para calcular los productos de ellas con la frecuencia absoluta y posteriormente hallar la sumatoria de dichos productos y finalmente promediar los valores obtenidos. Para ello, se agrega una nueva columna ($x_i * f_i$) a nuestra tabla:

Edad (x)	Marca de Clase (x_i)	Frecuencia Absoluta (f_i)	Frecuencia Absoluta Acumulada (F_i)	($x_i * f_i$)
13 – 15	14	4	4	56
15 – 17	16	9	13	144
17 – 19	18	3	16	54
19 – 21	20	3	19	60
21 – 23	22	1	20	22
Total		20		$\Sigma = 336$

*Recuerda que la **marca de clase** es la semisuma de los límites de dicho intervalo, para este ejercicio, note que la primera marca de clase es 14 ya que $x_i = \frac{13+15}{2} = \frac{28}{2} = 14$.

$\bar{x} = \frac{\Sigma x_i * f_i}{n}$	<p>Donde:</p> <ul style="list-style-type: none"> Σ = Sumatoria x_i = Marca de clase f_i = Frecuencia absoluta n = Número total de datos
--	--

Sustitución:

$$\bar{x} = \frac{336}{20} = 16.8 \text{ Años}$$

Mediana

La mediana es una medida de posición, entonces en los datos agrupados es necesario hallar la posición correcta del dato de en medio antes de aplicar la fórmula correspondiente, para ello primero utilizamos las fórmulas utilizadas en los datos no agrupados dependiendo si nuestro conjunto de datos es par o impar para encontrar la posición de la clase a utilizar.

Muestra Impar	Muestra Par
$\tilde{x} = \frac{n + 1}{2}$	$\tilde{x} = \frac{n}{2}$
Donde: $n = \text{Número total de datos}$	

En este caso nuestra muestra es par, pues $n = 20$:

$$\tilde{x} = \frac{n}{2} = \frac{20}{2} = 10$$

Nuestro primer resultado es 10, se procede a encontrar el dato número 10 en la columna de Frecuencia Absoluta acumulada de nuestra tabla:

Edad (x)	Marca de Clase (x_i)	Frecuencia Absoluta (f_i)	Frecuencia Absoluta Acumulada (F_i)	($x_i * f_i$)
13 – 15	14	4	4	56
15 – 17	16	9	13	144
17 – 19	18	3	16	54
19 – 21	20	3	19	60
21 – 23	22	1	20	22
	Total	20		$\Sigma = 336$

Como podemos ver, el número **10** no se encuentra, por lo que buscamos el número más cercano a este, sin que le falten unidades, en este caso es **13** quien nos indica la clase con la que se trabajará la aplicación de la fórmula.

Nota: En caso de encontrar la posición exacta en la frecuencia absoluta acumulada entonces el valor de la mediana es igual al límite superior del intervalo.

Fórmula para calcular la mediana de datos no agrupados:

$$\tilde{x} = L_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} * A_i$$

Donde:

$n = \text{Número total de datos}$

$L_i = \text{Limite inferior de la clase de la mediana}$

$F_{i-1} = \text{Frecuencia absoluta acumulada anterior a la clase de la mediana}$

$f_i = \text{Frecuencia absoluta de la clase de la mediana}$

$A_i = \text{Amplitud del intervalo}$

Sustitución de los datos

Datos:

$$n = 20$$

$$L_i = 15$$

$$F_i - 1 = 4$$

$$f_i = 9$$

$$A_i = 2$$

$$\tilde{x} = 15 + \frac{\frac{20}{2} - 4}{9} * 2$$

$$\tilde{x} = 15 + \frac{10 - 4}{9} * 2$$

$$\tilde{x} = 15 + \frac{6}{9} * 2$$

$$\tilde{x} = 15 + \frac{12}{9}$$

$$\tilde{x} = 16.33 \text{ Años}$$

Nótese que el valor de la mediana debe estar entre los valores del intervalo que, en este caso, es 15-17.

Moda

Recordemos que la moda es el dato que más se repite dentro de la muestra, por lo que observamos nuestra tabla de frecuencias en la columna de frecuencia absoluta y buscamos el número mayor que en este caso es **9**, entonces este será la clase con la que trabajaremos en la fórmula correspondiente.

Es importante señalar que no siempre será la misma clase para la mediana y la moda, casi siempre se trabaja con clases diferentes en ambas medidas.

Edad (x)	Marca de Clase (x _i)	Frecuencia Absoluta (f _i)	Frecuencia Absoluta Acumulada (F _i)	(x _i * f _i)
13 – 15	14	4	4	56
15 – 17	16	9	13	144
17 – 19	18	3	16	54
19 – 21	20	3	19	60
21 – 23	22	1	20	22
	Total	20		$\Sigma = 336$

*También puedes apoyarte con algún software para la realización de las tablas.



"Educación que genera cambio"

Fórmula para calcular la moda de datos no agrupados:

$$\tilde{x} = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} * A_i$$

Donde:

L_i : Límite inferior de la clase con mayor frecuencia.

f_i : Frecuencia absoluta del intervalo con mayor frecuencia.

f_{i-1} : Frecuencia absoluta anterior al intervalo de mayor frecuencia.

f_{i+1} : Frecuencia absoluta posterior al intervalo de mayor frecuencia

A_i : Amplitud del intervalo

Sustitución:

$$\tilde{x} = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} * A_i$$

$$\tilde{x} = 15 + \frac{9 - 4}{(9 - 4) + (9 - 3)} * 2$$

$$\tilde{x} = 15 + \frac{5}{(5) + (6)} * 2$$

$$\tilde{x} = 15 + \frac{5}{11} * 2$$

$$\tilde{x} = 15 + \frac{10}{11}$$

$$\tilde{x} = 15 + 0.9$$

$$\tilde{x} = 15.9 \text{ Años}$$

Datos:

$L_i = 15$

$f_i = 9$

$f_{i-1} = 4$

$f_{i+1} = 3$

$A_i = 2$



"Educación que genera cambio"

De la información al conocimiento



LÍMITES ESTADÍSTICOS Y MEDIDAS DE POSICIÓN

Las medidas de posición son valores para los cuales una porción específica de la distribución queda en o debajo de ellos, también son aquellas medidas que permiten dividir los datos en dos partes iguales, llamada *mediana* y esta medida de posición está dada con respecto a los datos centrales porque se divide en dos partes (50%). También se puede dividir en cuatro partes iguales llamadas *cuartiles*, en diez partes iguales llamadas *deciles* y en *percentiles*, es decir, en 100 partes iguales.

Todas estas medidas permiten obtener valores específicos e información resumida sobre la variable u objeto de estudio.

Como ya se mencionó, las medidas de posición son: **Percentiles, Cuartiles y Deciles.**

Percentiles:

Es un valor tal que, a lo más, un $p\%$ de los datos tienen dicho valor o menos y, al menos, un $(100 - p)\%$ de los datos tienen este valor o más.

Para calcular el p -ésimo (punto) percentil de un conjunto de n datos, es importante que tengamos en cuenta los siguientes pasos:

- Ordenar los datos de manera ascendente.
- Calculemos un índice (i) a través de la fórmula $i = \frac{np}{100}$, siendo p el percentil de interés y n , la cantidad de datos.
- Decidir de acuerdo con uno de los dos casos:
 - ✓ Si el índice i no es entero, redondeamos al entero siguiente. Este valor aproximado de i indica la posición de p -ésimo percentil.
 - ✓ Si i es entero, el p -ésimo percentil es el promedio de los valores de los datos ubicados en las posiciones i y $i + 1$.

Cuartiles:

Son las medidas de posición relativa correspondiente a un conjunto ordenado de datos divididos en cuatro partes (iguales) y se define así:

- Q_1 = primer cuartil o 25-ésimo percentil.
- Q_2 = segundo cuartil o 50-ésimo percentil o mediana.
- Q_3 = tercer cuartil o 75-ésimo percentil.

Al igual que los cuartiles, los deciles son también casos particulares de los percentiles.



"Educación que genera cambio"

Deciles:

Son las medidas de posición relativa correspondiente a un conjunto de datos (ordenados ascendentemente) que está dividido en diez partes, de tal forma que cada parte contiene aproximadamente 10% de las medidas.

Hay nuevos deciles, denotados por D_1, D_2, \dots y D_9 . Si D_n es el n-ésimo decil, entonces cada punto decil, corresponde a un punto percentil. Por ejemplo, D_4 es el 40-ésimo punto percentil, D_7 es el 70-ésimo punto percentil, etc.

DATOS NO AGRUPADOS.

Ejemplo:

1. Calcula el 85-ésimo punto percentil y el 50-ésimo punto percentil de los siguientes datos que representan los salarios (en millones de pesos) de 12 empleados en una empresa:

2350	2450	2550	2380	2255	2210	2390	2630	2440	2825	2420	2380
------	------	------	------	------	------	------	------	------	------	------	------

Solución:

Ordenar los datos de manera ascendentes:

2210, 2255, 2350, 2380, 2380, 2390, 2420, 2440, 2450, 2550, 2630, 2825

- a) Calcula el 85-ésimo punto percentil

Calculamos el índice i , $i = \frac{85 \cdot 12}{100} = \frac{1020}{100} = 10.2$, no es entero se redondea al entero siguiente: 11.

Por lo tanto, el lugar 85-ésimo punto percentil es el lugar 11, en los datos ordenados el 85% de los empleados de la empresa ganan \$ 2630.

- b) Calcula el 50-ésimo punto percentil.

Calculamos el índice i , $i = \frac{50 \cdot 12}{100} = \frac{600}{100} = 6$

Por lo tanto, el lugar 50-ésimo punto percentil es el promedio de los valores de sexto (\$2390) y séptimo (\$2420), $= \frac{2390+2420}{2} = \frac{4810}{2} = 2405$

El 50% de los empleados tiene un salario menor o igual a \$ 2405

2. Calcule los cuartiles de la tabla anterior.

Q_2 = coincide con la mediana, Q_2 es \$2450

$$Q_1 = \frac{2350 + 2380}{2} = 2365 \qquad Q_3 = \frac{2450 + 2550}{2} = 2500$$

El primer valor del cuartil indica que el 25% de los empleados gana al menos \$2365
 El segundo valor del cuartil indica que el 50% de los empleados gana al menos \$2405
 El tercer valor del cuartil indica que el 75% de los empleados gana al menos \$ 2500

3. Calcule el decil número 3

D_3 sería el 30-ésimo punto percentil, que corresponde \$2380

PARA DATOS AGRUPADOS

Percentiles:

$$p_k = L_i + \frac{\left(\frac{kn'}{100} - F_{i-1}\right)}{f_i} * C_i$$

Fórmula:

L_i = Límite Inferior del intervalo que contiene al percentil

F_{i-1} = Frecuencia acumulada en la clase anterior k-ésima

f_i = Frecuencia en la clase que contiene al percentil

c = Tamaño del intervalo de clase.

k = 1%, 2%, 3%, ..., 97%, 98%, 99% percentiles

1. Hallar el valor del percentil 60 y 80 del siguiente conjunto de estaturas de un grupo de estudiantes de un colegio.

Altura (cm)	Frecuencia f_i	Frecuencia Acumulada F_i
152 – 160	5	5
160 – 168	18	23
168 – 176	42	65
176 – 184	27	92
184 – 192	8	100
	N= 100	

a) Percentil 60

$\frac{kn}{100} = \frac{60 \cdot 100}{100} = 60$, la posición del percentil es 60 que se encuentra en el intervalo 168 y 176.

$$p_k = L_i + \frac{\left(\frac{kn'}{100} - F_{i-1}\right)}{f_i} * C_i$$

$$p_{60} = 168 + \frac{\left(\frac{60 * 100}{100} - 23\right)}{42} * 8$$

$$p_{60} = 175.05 \text{ cm}$$

El 60% de los estudiantes tienen una estatura menor o igual a 175.05 cm

b) Percentil 80

$\frac{kn}{100} = \frac{80 \cdot 100}{100} = 80$, la posición del percentil es 80 y se encuentra en el intervalo 176 y 184.

$$p_k = L_i + \frac{\left(\frac{kn'}{100} - F_{i-1}\right)}{f_i} * C_i$$

$$p_{80} = 176 + \frac{\left(\frac{80 * 100}{100} - 65\right)}{27} * 8$$

$$p_{80} = 180.4 \text{ cm}$$

El 80% de los estudiantes tienen una estatura menor o igual a 180.4 cm

Cuartiles:

$$Q_k = L_i + \frac{\left(\frac{kN}{4} - F_{i-1}\right)}{f_i} * a_i \quad k = 1, 2, 3$$

Fórmula:

L_i = Límite Inferior de la clase donde se encuentra el cuartil.

F_{i-1} = Frecuencia acumulada anterior a la clase del cuartil.

f_i = Frecuencia en la clase que contiene al cuartil

A_i = Tamaño del intervalo de clase.

N = es la suma de las secuencias absolutas



"Educación que genera cambio"

1. Calcular los cuartiles del siguiente conjunto de pesos de un grupo de estudiantes de un colegio.

Pesos (kg)	Frecuencia f_i	Frecuencia Acumulada F_i
50 – 60	8	8
60 – 70	10	18
70 – 80	16	34
80 – 90	14	48
90 – 100	12	60
	N= 60	

a) Cuartil Q_1

$\frac{kn}{4} = \frac{1 \cdot 60}{4} = 15$, se encuentra en el intervalo 60 y 70

$$Q_k = L_i + \frac{\left(\frac{kN}{4} - F_{i-1}\right)}{f_i} * a_i$$

$$Q_1 = 60 + \frac{\left(\frac{60 * 1}{4} - 8\right)}{10} * 10$$

$$Q_1 = 67 \text{ kg}$$

b) Q_2

$\frac{kn}{4} = \frac{2 \cdot 60}{4} = 30$, se encuentra en el intervalo 70 y 80

$$Q_k = L_i + \frac{\left(\frac{kN}{4} - F_{i-1}\right)}{f_i} * a_i$$

$$Q_2 = 70 + \frac{\left(\frac{60 * 2}{4} - 18\right)}{16} * 10$$

$$Q_2 = 77.5 \text{ kg}$$

c) Q_3

$\frac{kn}{4} = \frac{3 \cdot 60}{4} = 45$, se encuentra en el intervalo 80 y 90

$$Q_k = L_i + \frac{\left(\frac{kN}{4} - F_{i-1}\right)}{f_i} * a_i$$

$$Q_3 = 80 + \frac{\left(\frac{60 * 3}{4} - 34\right)}{14} * 10$$

$$Q_3 = 87.85 \text{ kg}$$

Deciles:

$$D_k = L_i + \frac{\left(\frac{kN}{10} - F_{i-1}\right)}{f_i} * C_i$$

Fórmula:

L_i = Límite Inferior de la clase donde pertenece el decil.

$K = 1, 2, 3$

F_{i-1} = Frecuencia acumulada anterior al decil.

f_i = Frecuencia absoluta del intervalo al que pertenece el decil.

C_i = Tamaño del intervalo de clase.

N = Número total de datos

- Hallar el valor del decil 3 y 8 del siguiente conjunto de estaturas de un grupo de estudiantes de un colegio.

Altura (cm)	Frecuencia f_i	Frecuencia Acumulada F_i
152 – 160	5	5
160 – 168	18	23
168 – 176	42	65
176 – 184	27	92
184 – 192	8	100
	$N = 100$	

a) Decil 3:

$\frac{kn}{10} = \frac{3 \cdot 100}{10} = 30$, se encuentra en el intervalo 168 y 176

$$D_k = L_i + \frac{\left(\frac{kN}{10} - F_{i-1}\right)}{f_i} * C_i$$

$$D_3 = 168 + \frac{\left(\frac{3 * 100}{10} - 23\right)}{42} * 8$$

$$D_3 = 169.28 \text{ cm}$$

b) Decil 8:

$\frac{kn}{10} = \frac{8 \cdot 100}{10} = 80$, se encuentra en el intervalo 176 y 184.

$$D_k = L_i + \frac{\left(\frac{kN}{10} - F_{i-1}\right)}{f_i} * C_i$$

$$D_8 = 176 + \frac{\left(\frac{8 * 100}{10} - 65\right)}{27} * 8$$

$$D_8 = 180.4 \text{ cm}$$



TABASCO

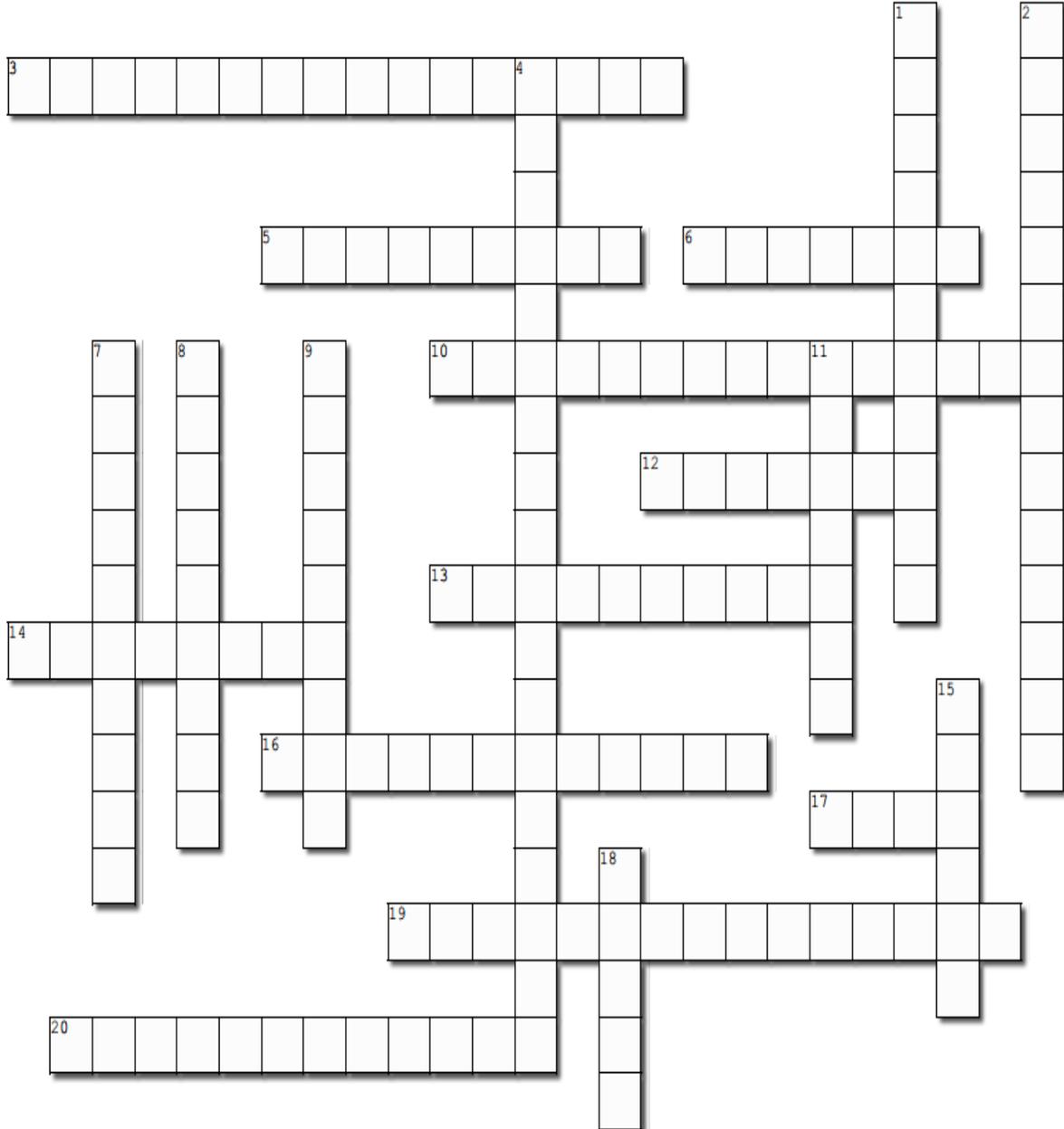
"Educación que genera cambio"



COBATAB

COLEGIO DE BACHILLERES DE TABASCO

Competencias a desarrollar		TAREA No. 4	CRUCIGRAMA: Medidas Estadísticas (Medidas centrales, Cuartiles, Deciles y Percentiles)
CG 5.1 CG 8.3	CDBM 1 CDBM 2 CDBM 3 CDBM 4		INSTRUCCIONES: Intégrate en binas y con ayuda de lo estudiado en las Lecturas 1 y 2, resuelvan el crucigrama sobre las Medidas Estadísticas



Horizontal:

3.- Se le conoce como medidas o parámetros que se sitúan hacia el centro de un conjunto de valores, con el objetivo de analizar un análisis profundo de los resultados

5.- Como parte de una campaña de salud, los estudiantes de sexto semestre se pesaron y los datos en kilos se presentan a continuación:

72,65,66,56,59,63,61,70,61,82,76,73,77,70,69,68.

¿Cuál es el peso máximo que obtuvieron la mitad de los estudiantes?

6.- Son 9 valores D1, D2; D3; D4; D5; D6; D7; D8 y D9 que dividen a un conjunto de datos en 10 partes iguales.

10.- Es el valor promedio de las observaciones de una muestra y se obtiene a partir de la suma de todos sus valores entre el total de ellas.

12.- Es el número de repeticiones máximas dentro de un conjunto de observaciones se puede presentar en dos valores.

13.- Es el número de veces en el que un dato o valor se repite dentro de un conjunto de datos

14.- Es la diferencia entre el límite superior y el límite inferior. Es la misma para cada clase.

16.- Es el valor medio de cada intervalo.

17.- Dentro de un conjunto de observaciones, es el valor que se repite con mayor frecuencia.

19.- Es el límite inferior y superior de un intervalo.

20.- Siete jóvenes compiten en una carrera de ciclismo de 300 metros. Los siguientes datos representan los tiempos, en segundos, obtenidos al final de la competencia.

36 42 45 39 33 36 48

¿Cuál fue el tiempo promedio que se necesitó para realizar el recorrido?

¿Cuál fue el tiempo promedio que se necesitó para realizar el recorrido?

Vertical:

1.- Son 99 valores que dividen a un conjunto de datos en 100 partes iguales.

2.- El doctor Manuel ha pedido a su secretaria el informe de las diversas enfermedades que se atendieron en la semana. Anticipadamente su secretaria registra el control de pacientes en la siguiente tabla.

Día	Paciente	Enfermedad
Lunes	Sra. Lolita Pérez	Resfriado común
	Sr. Pedro García	Infección del oído
Martes	Sr. Guillermo Rodríguez	Sinusitis
Miércoles	Pablo Jiménez	Dolor de garganta
Jueves	Berenice Remellón	Resfriado común
	Sra. Clara Hernández	Infección urinaria
Viernes	Sra. Karla Estrada	Resfriado común
	Sr. Juan Padilla	Resfriado común
Sábado	Sra. Karina Ramos	Infección del oído
	Ana Martínez	Bronquitis aguda

¿Cuál ha sido la enfermedad que representa la moda en este conjunto de datos?

4.- Es una ordenación de datos o valores estadísticos en forma de tabla.

7.- Dada la siguiente tabla, halla el percentil 80% de los gastos diarios en periódicos estará en intervalo 5

8.- Son 3 valores Q1; Q2 y Q3 que dividen a los datos en cuatro partes iguales.

9.- Un reporte de laboratorio indica el número de pacientes que fueron atendidos en los primeros 100 días del año, de las solicitudes enviadas por una clínica para que sus pacientes sean atendidos y se les realizaran estudios de glucosa. Determina el valor del primer cuartil.

Intervalos	Promedio de días X_i	Número de paciente f_i	Frecuencia acumulada F_i
1 a 10 días	5	5	5
10 a 20 días	14.5	6	11
20 a 30 días	24.5	8	19
30 a 40 días	34.5	8	27
40 a 50 días	44.5	4	31
50 a 60 días	54.5	5	36
60 a 70 días	64.5	7	43
70 a 80 días	74.5	8	51
80 a 90 días	84.5	4	55
90 a 100 días	94.5	8	63

11.- Es aquel valor que está en el lugar central dentro de un conjunto de datos, dividiéndolo en partes iguales.

15.- Es la agrupación de datos en intervalos.

18.- Establece el grado de simetría o asimetría de una representación gráfica de las medidas de centralización.



"Educación que genera cambio"

PYEI-B2-LC003 Lista de Cotejo para evaluar Tarea 04 Crucigrama

INSTRUMENTO DE EVALUACION

LISTA DE COTEJO

Actividad: CRUCIGRAMA "Medidas Estadísticas (Medidas centrales, Cuartiles, Deciles y Percentiles)"

DATOS GENERALES

Nombre(s) del alumno(s)	Matricula(s)
Producto:	Fecha
Materia:	Periodo
Nombre del docente	Firma del docente

VALOR DEL REACTIVO	CARACTERISTICAS A CUMPLIR	VALOR OBTENIDO		CALIF	OBSERVACIONES Y/O SUGERENCIAS DE MEJORA
		SI	NO		
1	Entregan la actividad en tiempo y forma establecidos.				
1	Entregan la actividad con limpieza y orden.				
1	Se integran con facilidad al equipo para el trabajo colaborativo.				
1	Respetan las opiniones e ideas de su compañero.				
1	Expresan sus ideas a través de diversos lenguajes (común, metodológico, etc.)				
1	Contribuyen con acciones para la resolución de problemas y reactivos del crucigrama.				
1	Comprende el tema de Medidas estadísticas, deciles, cuartiles, percentiles y los aplica correctamente a cada uno de los planteamientos dados.				
1	Identifican y relacionan con facilidad cada uno de los conceptos planteados.				
2	Utiliza acertadamente la metodología de operaciones necesarias para resolver los ejercicios propuestos.				
10	CALIFICACION				

Realimentación:

Logros:	Aspectos de mejora:

Firma del evaluador: _____



"Educación que genera cambio"



MEDIDAS DE TENDENCIA CENTRAL Y DE DISPERSIÓN PARA DATOS SIMPLES

Datos no agrupados.

Cuando los datos se presentan en una tabla de distribución de e frecuencias para datos no agrupados, es preciso modificar el procedimiento para el cálculo de las medidas de tendencia central.

Para guiarnos en este desarrollo, realizaremos un ejemplo guiado, planteando el siguiente caso:

En la cooperativa del plantel 06 se venden diariamente muchos productos, el encargado de ella se ha propuesto investigar cual es la cantidad de dinero que se gasta un estudiante en sus alimentos en el plantel, para ello realizo una encuesta que los estudiantes de probabilidad y estadística 1 le han ayudado a presentar como una tabla de distribución de frecuencias para datos no agrupados.

x_i	f_i	f_a	f_r	f_{ra}	$f_{r\%}$
35	6	6	0.17	0.17	17%
40	9	15	0.26	0.43	26%
45	7	22	0.20	0.63	22%
50	8	30	0.23	0.86	23%
60	5	35	0.14	1.00	14%
Σ	35		1		100%

El encargado de la cooperativa desea emplear esta información para diseñar un menú acorde a las necesidades de la población. Sin embargo, no sabe interpretar la información que se obtuvo.

Emplea las medidas de tendencia central para ofrecerle información que le permita realizar un menú adecuado.

Media:

Si se tiene un conjunto de datos, con valores numéricos denotados como X_i y estos se presentan a su vez con distintos valores de frecuencia f_i , entonces la fórmula para calcular la media aritmética sería

$$\bar{x} = \frac{\sum_1^n x_i f_i}{n}$$

Es decir, se tiene que realizar la suma de todos los productos que se obtiene a multiplicar cada valor por su respectiva media. Para hacerlo podemos emplear una columna adicional a nuestra tabla de distribución de frecuencias.

x_i	f_i	$x_i * f_i$	f_a	f_r	f_{ra}	$f_{r\%}$
35	6	210	6	0.17	0.17	17%
40	9	360	15	0.26	0.43	26%
45	7	315	22	0.20	0.63	22%
50	8	400	30	0.23	0.86	23%
60	5	300	35	0.14	1.00	14%
Σ	35	1585		1		100%

Al sumar todos los elementos de la columna tendríamos el valor correspondiente a la parte superior de la fórmula, la parte inferior corresponde al tamaño de la muestra, el cual lo podemos obtener sumando todas las frecuencias absolutas, de tal manera que podríamos sustituir ambos elementos en la fórmula y obtendríamos:

$$\bar{x} = \frac{\sum_1^n x_i f_i}{n} \rightarrow \bar{x} = \frac{1585}{35} = 45.28$$

De esta manera se obtiene la media aritmética para una tabla de frecuencias con datos no agrupados.

Mediana:

Calcular la mediana se base en el ordenamiento de los datos del conjunto, cuando se presentan estos en una tabla de frecuencias para datos no agrupados, se entiende que las clases tienen un ordenamiento ascendente (de menor a mayor).

Para entender esta parte es necesario comprender la frecuencia absoluta como un conjunto de posiciones, cada frecuencia absoluta tiene sus propias posiciones, por ejemplo:

x_i	f_i	f_a	f_r	f_{ra}	$f_{r\%}$
35	6	6	0.17	0.17	17%
40	9	15	0.26	0.43	26%
45	7	22	0.20	0.63	22%
50	8	30	0.23	0.86	23%
60	5	35	0.14	1.00	14%
Σ	35		1		100%

$f_a = 6 \quad i = \{1, 2, 3, 4, 5, 6\}$
 $f_a = 15 \quad i = \{7, 8, 9, 10, 11, 12, 13, 14, 15\}$
 $f_a = 22 \quad i = \{15, 16, 17, 18, 19, 20, 21, 22\}$
 $f_a = 30 \quad i = \{23, 24, 25, 26, 27, 28, 29, 30\}$
 $f_a = 35 \quad i = \{31, 32, 33, 34, 35\}$

Para encontrar la mediana es necesario determinar el procedimiento a seguir, para ello es necesario corroborar si el valor del tamaño del conjunto de datos es par o impar.

Cuando n es impar la posición se obtiene de la siguiente manera.

$$p = \frac{n}{2} + 0.5$$

Cuando n es par se obtendrán dos posiciones de las siguientes maneras.

$$p_1 = \frac{n}{2} \quad p_2 = \frac{n}{2} + 1$$

Estas posiciones se emplearán para encontrar la mediana mas no son el valor de ella, para encontrar el valor de esta es necesario usar la columna de frecuencia absoluta (f_a) donde están contenidas, y emplear el valor X_i que le corresponde.

Por ejemplo:

x_i	f_i	f_a
6	4	4
7	6	10
8	2	12
9	3	15
10	5	20
Σ	20	

Si $p = 13 \quad x_{13} = 9$

Si $p = 3 \quad x_3 = 6$



"Educación que genera cambio"

Quando n es impar la Mediana se obtiene de la siguiente manera.

$$\hat{x} = x_p$$

Por ejemplo:

x_i	f_i	f_a
6	4	4
7	6	10
8	2	12
9	3	15
10	5	20
Σ	25	

$$p = \frac{25}{2} + 0.5 = 13$$

$$\hat{x} = x_{13} = 9$$

La mediana es 9

Quando n es par La mediana se obtiene de la siguiente manera .

$$\hat{x} = \frac{x_{p_1} + x_{p_2}}{2}$$

Por ejemplo:

x_i	f_i	f_a
6	4	4
7	6	10
8	2	12
9	3	15
10	5	20
Σ	20	

$$p_1 = \frac{20}{2} = 10 \quad p_2 = \frac{20}{2} + 1 = 11$$

$$\hat{x} = \frac{x_{10} + x_{11}}{2}$$

$$\hat{x} = \frac{7 + 8}{2} = \frac{15}{2} = 7.5$$

La mediana es 7.5

Para el problema que estamos desarrollando usaríamos:

$$p = \frac{n}{2} + 0.5 \rightarrow p = \frac{35}{2} + 0.5 \rightarrow p = 17.5 + 0.5$$

$$p = 18$$

La mediana se calcula con la posición:

$$\hat{x} = x_p$$

$$\hat{x} = x_{18}$$

$$\hat{x} = 45$$

Moda:

La moda está dada por el dato que más se repite, al encontrarse los datos en una tabla de frecuencias para datos no agrupados, es fácil identificar la moda como el valor de la clase con mayor frecuencia absoluta.

x_i	f_i	f_a	f_r	f_{ra}	$f_{r\%}$
35	6	6	0.17	0.17	17%
40	9	15	0.26	0.43	26%
45	7	22	0.20	0.63	22%
50	8	30	0.23	0.86	23%
60	5	35	0.14	1.00	14%
$\sum x_i$	35		1		100%

Primero buscamos en la columna de las frecuencias absolutas f_i y ubicamos el valor más alto, en este ejemplo sería el 9, posteriormente seleccionamos el valor de la clase x_i y ese es el valor de la moda.

$$\tilde{x} = 40$$

Sesgo:

El sesgo es una unidad que representa la asimetría del polígono de frecuencias de una distribución, esto es la gráfica de una tabla de frecuencias para datos no agrupados, para calcularlo hay varias maneras, la que usaremos sería la dada por la siguiente fórmula.

$$Sesgo = \frac{Media - Moda}{desviacion\ estandar}$$

$$Sesgo = \frac{\bar{x} - \tilde{x}}{s}$$

Existen 3 interpretaciones para el sesgo dependiendo de su signo y valor.

- Si el sesgo es positivo, se tiene una asimetría hacia la derecha.
- Si el sesgo es negativo, se tiene una asimetría hacia la izquierda.
- Si el sesgo es 0, entonces no hay asimetría.

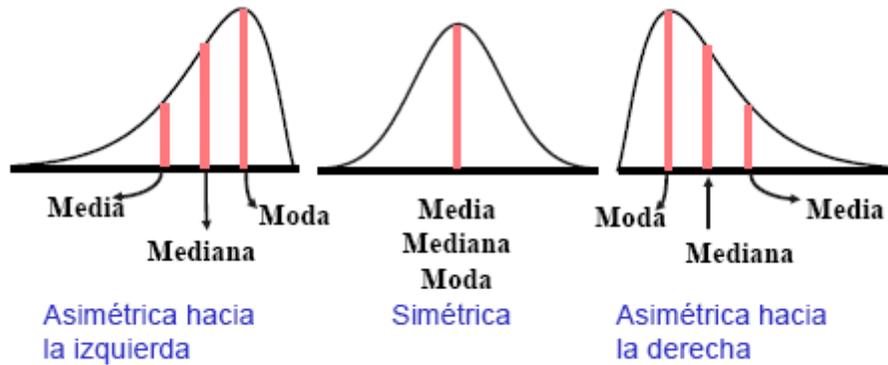


Ilustración tomada de: http://1.bp.blogspot.com/_mA-bypZU_yA/ScfGYolZY3I/AAAAAAAAA4/g4_2B4EJ4c/s1600/image028.gif

Para el ejemplo que estamos desarrollando, nos hace falta obtener la desviación estándar la cual aprenderemos más adelante en la lectura, para calcular el sesgo usaremos solo su valor el cual es 7.83. Ahora sustituimos los datos en la fórmula:

$$\begin{aligned} \text{Sesgo} &= \frac{45.28 - 40}{7.83} \\ \text{Sesgo} &= \frac{5.28}{7.83} \\ \text{Sesgo} &= 0.67 \end{aligned}$$

Por lo cual podemos deducir que la distribución tiene una asimetría a la derecha.

Rango

El rango de un conjunto de números es la diferencia entre el número mayor y el número menor del conjunto.

$$\text{Rango} = \text{Dato mayor} - \text{Dato menor}$$

Para obtener los datos menor y mayor nos referimos a la tabla de distribución de frecuencias para datos no agrupados de nuestro ejemplo contextualizado.

x_i	f_i	f_a	f_r	f_{ra}	$f_{r\%}$
35	6	6	0.17	0.17	17%
40	9	15	0.26	0.43	26%
45	7	22	0.20	0.63	22%
50	8	30	0.23	0.86	23%
60	5	35	0.14	1.00	14%
$\sum x_i$	35		1		100%

El dato menor es la clase que posea el menor valor y el dato mayor es el valor de la clase con mayor valor. Para nuestro ejemplo la sustitución en la fórmula sería:

$$\text{Rango} = 60 - 35$$

$$\text{Rango} = 25$$

Desviación media

La desviación media es el promedio de las desviaciones (separaciones) de cada elemento de la distribución con respecto a la media aritmética de la distribución. Su fórmula para tablas de frecuencias de datos no agrupados sería:

$$Dm = \frac{\sum_1^n (|x_i - \bar{x}|) f_i}{n}$$

Para calcularla nos enfocaremos en las columnas x_i y f_i , y agregaremos dos columnas auxiliares,

$|x_i - \bar{x}|$ La cual se obtiene al aplicar el valor absoluto al resultado de la resta de cada clase menos el valor de la media.

$f_i |x_i - \bar{x}|$ La cual se obtiene al multiplicar la frecuencia absoluta f_i por la columna $|x_i - \bar{x}|$

En nuestro ejemplo contextualizado la tabla quedaría:

x_i	f_i	$ x_i - \bar{x} $	$f_i x_i - \bar{x} $	f_a	f_r	f_{ra}	$f_{r\%}$
35	6	10.28	61.68	6	0.17	0.17	17%
40	9	5.28	47.52	15	0.26	0.43	26%
45	7	0.28	1.96	22	0.20	0.63	22%
50	8	4.72	37.76	30	0.23	0.86	23%
60	5	14.72	73.6	35	0.14	1.00	14%
Σ	35		222.52		1		100%

Ahora realizamos la suma de todos los elementos de la columna $f_i|x_i - \bar{x}|$ y sustituimos los datos en la fórmula:

$$Dm = \frac{222.52}{35}$$

$$Dm = 6.35$$

Varianza

La varianza determina que tanto se alejan los datos de la media, considerando el signo de sus desviaciones, para calcularla en tablas de frecuencias para datos no agrupados se emplea la siguiente formula:

$$s^2 = \frac{\sum_1^n f_i(x_i - \bar{x})^2}{n}$$

Para poder emplear la formula usaremos las columnas x_i y f_i , además de tres columnas auxiliares:

$x_i - \bar{x}$ Es el resultado de restar al valor de la clase x_i el valor de la media \bar{x} .

$(x_i - \bar{x})^2$ Es el resultado de elevar al cuadrado la columna $x_i - \bar{x}$

$f_i(x_i - \bar{x})^2$ Es el resultado de multiplicar la frecuencia absoluta f_i , por la columna $(x_i - \bar{x})^2$



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

x_i	f_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$	f_a	f_r	f_{ra}	$f_r\%$
35	6	-10.28	105.6784	634.0704	6	0.17	0.17	17%
40	9	-5.28	27.8784	250.9056	15	0.26	0.43	26%
45	7	-0.28	0.0784	0.5488	22	0.20	0.63	22%
50	8	4.72	22.2784	178.2272	30	0.23	0.86	23%
60	5	14.72	216.6784	1083.392	35	0.14	1.00	14%
Σ	35			2147.14		1		100%

Después realizamos la suma de todos los elementos de la columna $f_i(x_i - \bar{x})^2$, con esos datos sustituyo los valores en la fórmula:

$$s^2 = \frac{2147.14}{35}$$

$$s^2 = 61.34$$

Desviación estándar

La desviación estándar se obtiene de la varianza, su fórmula es:

$$s = \sqrt{\frac{\sum_i^n f_i(x_i - \bar{x})^2}{n}} \quad \text{o} \quad s = \sqrt{s^2}$$

Al conocer la varianza solo la sustituimos en la segunda fórmula:

$$s = \sqrt{61.34}$$

$$s = 7.83$$



"Educación que genera cambio"

MEDIDAS DE TENDENCIA CENTRAL Y DE DISPERSIÓN PARA DATOS AGRUPADOS

Datos agrupados.

Cuando los datos se presentan en una tabla de distribución de e frecuencias para datos agrupados, es preciso modificar el procedimiento para el cálculo de las medidas de tendencia central y de dispersión.

Para guiarnos en este desarrollo, realizaremos un ejemplo guiado, planteando el siguiente caso:

En el plantel 06 del COBATAB la academia de orientación educativa realizo un estudio sobre los hábitos académicos de los estudiantes, el estudio se enfocó en las horas que estos dedican al estudio en una semana normal, los resultados se presentan en la siguiente tabla de frecuencias para datos agrupados.

<i>Li</i>	<i>Ls</i>	<i>Mc</i>	<i>f_i</i>	<i>f_a</i>	<i>f_r</i>	<i>f_{ra}</i>	<i>f_r%</i>
0	4	2	24	24	0.24	0.24	24%
4	8	6	31	55	0.31	0.55	31%
8	12	10	13	68	0.13	0.68	13%
12	16	14	21	89	0.21	0.89	21%
16	20	18	11	100	0.11	1	11%
			100				

Para realizar los cálculos solo emplearemos las primeras cinco columnas.

Media

El cálculo de la media sigue el mismo procedimiento que realizamos para las tablas de frecuencias para datos no agrupados, sin embargo, se cambia la clase (xi) por la marca de clase (mc) para realizar los cálculos, por ese motivo la fórmula de la media para tablas de frecuencias para datos agrupados quedaría de la siguiente manera:

$$\bar{x} = \frac{\sum_{i=1}^n f_i mc}{n}$$

Para poder emplearla, debemos agregar una columna donde multipliquemos la marca de clase(mc) de cada una de las filas por su correspondiente frecuencia absoluta (fi)



"Educación que genera cambio"

Li	Ls	Mc	f_i	$f_i \cdot Mc$	f_n
0	4	2	24	48	24
4	8	6	31	186	55
8	12	10	13	130	68
12	16	14	21	294	89
16	20	18	11	198	100
			100	856	

Al sumar todos los elementos de la columna tendríamos el valor correspondiente a la parte superior de la fórmula, la parte inferior corresponde al tamaño de la muestra, el cual lo podemos obtener sumando todas las frecuencias absolutas, de tal manera que podríamos sustituir ambos elementos en la fórmula y obtendríamos:

$$\bar{x} = \frac{856}{100} \rightarrow \bar{x} = 8.56$$

De esta manera se obtiene la media aritmética para una tabla de frecuencias con datos agrupados.

Mediana

La mediana se basa en las posiciones, por lo cual retomaremos lo aprendido en la lectura anterior sobre las posiciones contenidas en las frecuencias absolutas.

Para calcular la mediana en una tabla de frecuencias es necesario determinar el intervalo mediano, para ello debemos usar alguna de las siguientes fórmulas:

$$p = \frac{n}{2} \text{ si es par} \qquad p = \frac{n}{2} + 0.5 \text{ si es impar}$$

Una vez que se obtiene la posición se ubica el valor de la frecuencia acumulada donde está contenido, ese intervalo será el intervalo mediano.

En nuestro ejemplo, primero calculamos la posición p . al ser $n = 100$ es un número par por lo que se usaría la siguiente fórmula:

$$p = \frac{100}{2} \qquad p = 50$$

Ahora ubicamos el valor de la frecuencia absoluta donde este contenido y determinamos el intervalo mediano.



"Educación que genera cambio"

Li	Ls	Mc	f_i	f_a
0	4	2	24	24
4	8	6	31	55
8	12	10	13	68
12	16	14	21	89
16	20	18	11	100
				100

Conociendo el intervalo mediano obtendremos varios datos a partir de él y podemos emplear la siguiente fórmula para encontrar la mediana.

$$\hat{x} = Li_{med} + \left(\frac{\frac{n}{2} - f_{a\ ant}}{f_{i\ med}} \right) Ac$$

Donde:

- Li_{med} = limite inferior del intervalo mediano
- n = numero de elementos de la muestra
- $f_{a\ ant}$ = frecuencia acumulada del intervalo anterior al intervalo mediano
- $f_{i\ med}$ = frecuencia absoluta del intervalo mediano
- Ac = amplitud de clase ($Ls - Li$)

Si sustituimos los datos en la formula obtenemos;

$$\hat{x} = 4 + \left(\frac{\frac{100}{2} - 24}{55} \right) 4$$

$$\hat{x} = 4 + \left(\frac{50 - 24}{55} \right) 4$$

$$\hat{x} = 4 + \left(\frac{26}{55} \right) 4$$

$$\hat{x} = 4 + (0,47) 4$$

$$\hat{x} = 4 + 1.88$$

$$\hat{x} = 5.88$$

De esta manera se calcula la mediana para tablas de frecuencias para datos agrupados.

Moda

La moda está dada por el dato que más se repite, al encontrarse los datos en una tabla de frecuencias para datos agrupados, es necesario determinar el intervalo modal primero, esto es muy fácil ya que es aquel o aquellos que posean el valor más alto de frecuencia absoluta (f_i).

Li	Ls	Mc	f_i	f_a
0	4	2	24	24
4	8	6	31	55
8	12	10	13	68
12	16	14	21	89
16	20	18	11	100
100				

Una vez identificamos el intervalo moda, tenemos que considerar las frecuencias absolutas anterior y posterior al intervalo modal. Y encontrar la diferencia respecto a la frecuencia absoluta del intervalo modal para después emplear la fórmula de la moda que se presenta:

$$\tilde{x} = Li_{mod} + \left(\frac{d_a}{d_a + d_p} \right) Ac$$

Donde:

- Li_{mod} = límite inferior del intervalo modal
- d_a = Diferencia respecto a la frecuencia anterior = $f_{mod} - f_{mod-1}$
- d_p = Diferencia respecto a la frecuencia posterior = $f_{mod} - f_{mod+1}$
- Ac = amplitud de clase ($Ls - Li$)

Primero calculamos las diferencias:

$$d_a = 31 - 24 = 7$$

$$d_p = 31 - 13 = 18$$

Sustituimos los datos en la fórmula:

$$\tilde{x} = 4 + \left(\frac{7}{7+18} \right) 4$$

$$\tilde{x} = 4 + \left(\frac{7}{25} \right) 4$$

$$\tilde{x} = 4 + (0.28) 4$$

$$\tilde{x} = 4 + 1.12$$

$$\tilde{x} = 5.12$$

Sesgo:

La definición para el sesgo es la misma para datos agrupados o no agrupados, por lo que esta vez se omitirá y sólo se retoma su fórmula:

$$\text{Sesgo} = \frac{\text{Media} - \text{Moda}}{\text{desviación estandar}}$$

$$\text{Sesgo} = \frac{\bar{x} - \tilde{x}}{s}$$

Recuerda:

- Si el sesgo es positivo, se tiene una asimetría hacia la derecha.
- Si el sesgo es negativo, se tiene una asimetría hacia la izquierda.
- Si el sesgo es 0, entonces no hay asimetría.

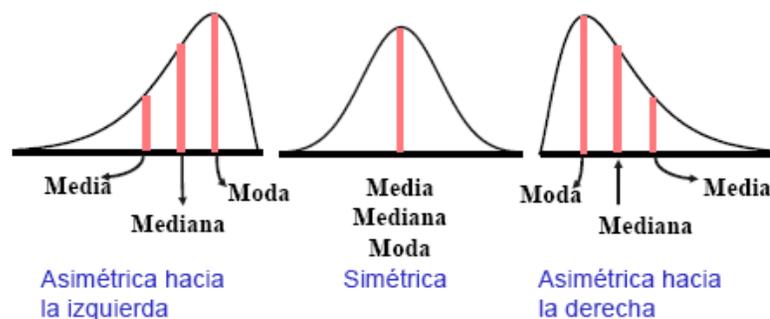


Ilustración tomada de: http://1.bp.blogspot.com/_mA-bypZU_yA/ScfGYolZY3I/AAAAAAAAA4/g4_2B4EJ4c/s1600/image028.gif

Para el ejemplo que estamos desarrollando, nos hace falta obtener la desviación estándar la cual aprenderemos a calcular más adelante en la lectura, para calcular el sesgo usaremos solo su valor el cual es 5.35. Ahora sustituimos los datos en la fórmula:

$$\text{Sesgo} = \frac{8.56 - 5.12}{5.35}$$

$$\text{Sesgo} = \frac{3.44}{5.35}$$

$$\text{Sesgo} = 0.64$$

Por lo cual podemos deducir que la distribución tiene una asimetría a la derecha.

Rango

El rango de un conjunto de números es la diferencia entre el número mayor y el número menor del conjunto.

$$Rango = \text{Dato mayor} - \text{Dato menor}$$

Para obtener los datos menor y mayor nos referimos a la tabla de distribución de frecuencias para datos agrupados; el dato mayor será el valor del límite superior (Ls) del último intervalo, mientras que el dato menor será el límite inferior (Li) del primer intervalo.

<i>Li</i>	<i>Ls</i>	<i>Mc</i>	<i>f_i</i>	<i>f_a</i>
0	4	2	24	24
4	8	6	31	55
8	12	10	13	68
12	16	14	21	89
16	20	18	11	100
			100	

Para nuestro ejemplo la sustitución en la fórmula sería:

$$Rango = 20 - 0$$

$$Rango = 20$$

Desviación media

La desviación media es el promedio de las desviaciones (separaciones) de cada elemento de la distribución con respecto a la media aritmética de la distribución. Su fórmula para tablas de frecuencias de datos agrupados sería:

$$Dm = \frac{\sum_{i=1}^n (|mc - \bar{x}|) f_i}{n}$$

Para calcularla agregaremos dos columnas auxiliares a nuestra tabla de frecuencias:

$|mc - \bar{x}|$ La cual se obtiene al aplicar el valor absoluto al resultado de la resta de cada marca de clase menos el valor de la media para cada uno de los intervalos.

$f_i |mc - \bar{x}|$ La cual se obtiene al multiplicar la frecuencia absoluta f_i por la columna $|mc - \bar{x}|$

En nuestro ejemplo contextualizado la tabla quedaría:

Li	Ls	Mc	f_i	$ mc - \bar{x} $	$f_i mc - \bar{x} $	f_a
0	4	2	24	6.56	157.44	24
4	8	6	31	2.56	79.36	55
8	12	10	13	1.44	18.72	68
12	16	14	21	5.44	114.24	89
16	20	18	11	9.44	103.84	100
			100	473.6		

Ahora realizamos la suma de todos los elementos de la columna $f_i|mc - \bar{x}|$ y sustituimos los datos en la fórmula:

$$Dm = \frac{473.6}{100}$$

$$Dm = 4.73$$

Varianza

La varianza determina que tanto se alejan los datos de la media, considerando el signo de sus desviaciones, para calcularla en tablas de frecuencias para datos no agrupados se emplea la siguiente formula:

$$s^2 = \frac{\sum_1^n f_i (mc - \bar{x})^2}{n}$$

Para poder emplear la formula usaremos las columnas x_i y f_i , además de tres columnas auxiliares:

- $mc - \bar{x}$ Es el resultado de restar al valor de la clase x_i el valor de la media \bar{x} .
- $(mc - \bar{x})^2$ Es el resultado de elevar al cuadrado la columna $x_i - \bar{x}$
- $f_i(mc - \bar{x})^2$ Es el resultado de multiplicar la frecuencia absoluta f_i , por la columna $(x_i - \bar{x})^2$

Li	Ls	Mc	f_i	$mc - \bar{x}$	$(Mc - \bar{x})^2$	$f_i(Mc - \bar{x})^2$	f_a
0	4	2	24	-6.56	43.0336	1032.8064	24
4	8	6	31	-2.56	6.5536	203.1616	55
8	12	10	13	1.44	2.0736	26.9568	68
12	16	14	21	5.44	29.5936	621.4656	89
16	20	18	11	9.44	89.1136	980.2496	100
			100	2864.64			



"Educación que genera cambio"

Después realizamos la suma de todos los elementos de la columna $f_i(mc - \bar{x})^2$, con esos datos sustituyo los valores en la fórmula:

$$s^2 = \frac{2864.64}{100}$$

$$s^2 = 28.64$$

Desviación estándar

- La desviación estándar se obtiene de la varianza, su fórmula es:

$$s = \sqrt{\frac{\sum_1^n f_i(mc - \bar{x})^2}{n}} \quad \text{o} \quad s = \sqrt{s^2}$$

- Al conocer la varianza solo la sustituimos en la segunda fórmula:

$$s = \sqrt{28.64}$$

$$s = 5.35$$

Competencias a desarrollar		TAREA No. 5	PROBLEMARIO: Medidas de centralización y de dispersión
CG. 1.1	CDBM. 1		INSTRUCCIONES: Emplea los conocimientos desarrollados en clase para resolver las situaciones que se te presentan.
CG. 4.1	CDBM. 2		
CG. 5.1	CDBM. 3		
CG. 5.2	CDBM. 4		
CG. 5.6	CDBM. 5		
CG. 8.2	CDBM. 8		

1). En una zapatería, hicieron un estudio de los últimos 25 pares de zapatos vendidos, analizando las tallas de zapatos y su frecuencia de ventas, como resultado obtuvieron la siguiente tabla.

Talla	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa Acumulada	Porcentaje
5	6	0.24	0.24	24%	
6	4	0.16	0.40	40%	
7	7	0.28	0.68	68%	
9	5	0.20	0.88	88%	
12	3	0.12	1	100%	
Σ	25	1			

Para su análisis se te solicita realizar los cálculos de las medidas de tendencia central y las medidas de dispersión, rellenando la siguiente tabla.

Medidas de centralización		Mediadas de dispersión	
Media		Rango	
Mediana		Desviación media	
Moda		Varianza	
Sesgo		Desviación estándar	

Del análisis de tus resultados responde las siguientes preguntas:

1. ¿Cuál es la talla de zapatos más vendida?
2. ¿Cuál es la talla promedio de los clientes?
3. ¿Existe una variación considerable entre las tallas de los clientes?
4. ¿Qué tallas consideras que deben tener más inventario?

Problema 2. En un parque de diversiones, se realiza un estudio sobre los visitantes, dentro del cual se toma en consideración las edades de estos, como resultado se obtiene la siguiente tabla:

Li	Ls	Mc	f_i	f_a
1	13	7	6	6
13	25	19	14	20
25	37	31	7	27
37	49	43	10	37
49	61	55	13	50
		Σ	50	

Para su análisis se te solicita realizar los cálculos de las medidas de tendencia central y las medidas de dispersión, rellenando la siguiente tabla.

Medidas de centralización		Mediadas de dispersión	
Media		Rango	
Mediana		Desviación media	
Moda		Varianza	
Sesgo		Desviación estándar	

Del análisis de tus resultados responde las siguientes preguntas:

1. ¿Qué edad tienen la mayoría de los visitantes?
2. ¿Cuál es la edad promedio de los visitantes?
3. ¿Para qué rango de edades se deben crear más atracciones?
4. ¿Existe mucha variación en las edades de los visitantes?



"Educación que genera cambio"

PYEI-B2-LC004 Lista de Cotejo para evaluar Tarea 05 Problemario

INSTRUMENTO DE EVALUACION

LISTA DE COTEJO

Actividad: Problemario de medidas de centralización y dispersión.

DATOS GENERALES

Nombre(s) del alumno(s)	Matricula(s)
Producto:	Fecha
Materia:	Periodo
Nombre del docente	Firma del docente

VALOR DEL REACTIVO	CARACTERISTICAS A CUMPLIR	VALOR OBTENIDO		CALIF	OBSERVACIONES Y/O SUGERENCIAS DE MEJORA
		SI	NO		
1	Realiza el cálculo correcto de las medidas de centralización para datos no agrupados				
1	Realiza el cálculo correcto de las medidas de dispersión para datos no agrupados				
1	Analiza los resultados y responde los cuestionamientos del problema 1				
1	Realiza el cálculo correcto de las medidas de centralización para datos agrupados				
1	Realiza el cálculo correcto de las medidas de dispersión para datos agrupados				
1	Analiza los resultados y responde los cuestionamientos del problema 2				
1	Entrega en tiempo y forma				
7	CALIFICACION				

Realimentación:

Logros:	Aspectos de mejora:

Firma del evaluador: _____



"Educación que genera cambio"

PYEI-B3-PP02		CUESTIONARIO TIPO PLANEA BLOQUE III
CG. 5.4 CG. 5.5 CG. 9.5	CDBM. 1 CDBM. 2 CDBM. 6 CDBM. 7	Indicaciones Lee cuidadosamente cada reactivo e individualmente elige la opción de respuesta correcta

1.- La antigüedad, en años, de los automóviles que repararon la semana pasada en un taller mecánico, es la siguiente:

5, 6, 3, 6, 11, 7, 9, 10, 2, 4, 10, 6, 2, 1, 5, 5, 4, 3

Indica cuál es la moda, dentro del conjunto de datos.

- a) 5
- b) 6
- c) 5;6
- d) No hay moda

2.- Una empresa lleva un registro del kilometraje recorrido por sus vehículos. A continuación, se presentan los registros semanales de sus 20 camiones:

810, 450, 756, 789, 210, 657, 589, 488, 876, 689, 1450, 560, 469, 890, 987, 559, 788, 943, 447, 775

Calcula la mediana del kilometraje que recorren los vehículos.

- a) $Me=709.1$
- b) $Me=722.5$
- c) $Me=723.5$
- d) $Me=1445$

3.- En un plantel de COBATAB se llevó a cabo la votación de 6 candidatas a reina de la primavera. El total de estudiantes que votaron fue de 1240, de la siguiente manera:

- Isabel: 88 votos
- Carmen: 230 votos
- Lucía: 125 votos
- Rosa: 302 votos
- Stephanie: 150 votos
- Jenny: 345 votos
- Total, de Votos: 1240

Para ver el comportamiento de la votación, calcula la variación media que existe entre estas cantidades de votos.

- a) 85.67
- b) 106.33
- c) 138.33
- d) 206.67

4.- A un grupo de personas se les pregunto la edad, resultando los siguientes datos:

Edad	f_i
30-35	2
35-40	4
40-45	8
45-50	5
50-55	1

¿Cuál es la desviación media resultante?

- a) 42.25
- b) 845
- c) 3.85
- d) 77



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

5.- En un consultorio médico, un pediatra obtuvo la siguiente tabla sobre la consulta de 50 niños en un mes, de acuerdo con su edad:

Edades (Meses)	9	10	11	12	13	14	15
Frecuencia de consulta	1	4	9	16	11	8	1

¿Cuál es la desviación media de la asistencia al pediatra en el mes?

- a) 13.51
- b) 11
- c) 7.14
- d) 4.41

6.- La tía Fabiola repartió entre sus sobrinos cierta cantidad de caramelos, como se muestra en la siguiente tabla:

Sobrinos	Caramelos
Víctor	12
Vicente	11
Harol	11
José	10
Carlos	11

¿Cuál es la desviación media de la repartición de Caramelos?

- a) 11
- b) 0.4
- c) 0.6
- d) 0

7.- Los siguientes datos representan la duración en segundos de 8 espacios comerciales televisivos, que fueron elegidos al azar y transmitidos por la televisora local.

18, 25, 30, 20, 15, 25, 28, 15

Deduce el valor aproximado de la desviación estándar.

- a) 22
- b) 33.14
- c) 5.75
- d) Ninguna de las anteriores

8.- De acuerdo con los datos que se indican en la siguiente tabla; serie A y B

Serie A	Serie B
20.3	20.9
20.8	20.5
21.0	24.0
20.5	29.5
21.1	21.0
20.2	19.1
20.4	16.4

¿En cuál serie los datos están menos dispersos?

- a) Serie A
- b) Serie B
- c) A y B son iguales
- d) Serie A con 13.1



"Educación que genera cambio"

9.- La distribución de las puntuaciones en una escala de hostilidad, de 160 sujetos de una muestra, ha sido la siguiente:

X_i	n_i	N_i
0 a 10	8	8
10 a 20	22	30
20 a 30	32	62
30 a 40	44	106
40 a 50	28	134
50 a 60	20	154
60 a 70	6	160

¿Cuál es el valor del percentil 27?

- a) 45
- b) 23.125
- c) 24.125
- d) 43.2

10.- La siguiente tabla muestra a cinco alumnos de alto rendimiento, de una escuela que realizaron un salto de longitud:

Alumno	Longitud del salto (m)
Jesús	4.4
Carlos	5.2
Alfredo	4.5
Sergio	5.3
Arturo	4.6

¿Calcula la longitud media alcanzada por los atletas?

- a) 5.2
- b) 4.4
- c) 4.6
- d) 4.8



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

PYEI-B3-MA02 Mapa de aprendizaje para evaluar los Aprendizajes Esperados

COLEGIO DE BACHILLERES DE TABASCO

Probabilidad y estadística I

Nombre	
Autoevaluación	Bloque III Medidas estadísticas

Mapa de aprendizaje

1: Necesito ayuda

2: Puedo hacerlo solo

3: Puedo ayudar a otros

Aprendizajes esperados	Nivel			Que debo hacer para mejorar:
	1	2	3	
Examina de manera crítica información resumida, por medio de las medidas de tendencia central obtenidas en diferentes contextos.				
Emplea comparaciones entre la dispersión de datos que le permitan estimar su comportamiento producto de una investigación y los datos que arrojan las medidas estadísticas para la toma consciente de decisiones.				
Usa límites estadísticos para el análisis crítico y reflexivo de situaciones problema.				

Estudiante	Facilitador
-------------------	--------------------

Fecha



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

REFERENCIAS BLOQUE III

Colegio de Bachilleres del Estado de Sonora. (2009). *Probabilidad y Estadística 1*.
Hermosillo, Sonora: COBAES.

Garza Olvera, B. (2017). *Estadística y probabilidad*. Cd de México: Pearson.

Stephens, L. J., & Spiegel, M. R. (2009). *ESTADÍSTICA*. Cd de México: Mc graw hill.



BLOQUE IV:
Comportamiento de dos
variables

Propósito del bloque IV



Bloque IV

Examina el comportamiento de datos bivariados con modelos matemáticos en situaciones de su contexto para la toma responsable de decisiones.

Aprendizajes Esperados

Ejemplifica críticamente por medio de modelos matemáticos la relación existente entre dos variables, así como su representación tabular y gráfica, que le permita afirmar o negar hipótesis planteadas en situaciones de su entorno



TABASCO

"Educación que genera cambio"



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

Competencias

Genéricas

5. Desarrolla innovaciones y propone soluciones a problemas a partir de métodos establecidos

5.4 Construye hipótesis y diseña y aplica modelos para probar su validez.

5.5 sintetiza evidencias obtenidas mediante la experimentación para producir conclusiones y formular nuevas preguntas

9. Participa con una conciencia cívica y ética en la vida de su comunidad, región, México y el mundo.

9.5 actúa de manera propositiva frente a fenómenos de la sociedad y se mantiene informado.

Disciplinares

1. Construye e interpreta modelos matemáticos mediante la aplicación de procedimientos aritméticos, algebraicos geométricos y variacionales para la comprensión y análisis de situaciones reales hipotéticas o formales

2. Formula y resuelve problemas matemáticos aplicando diferentes enfoques

6. Cuantifica, representa y contrasta experimental o matemáticamente las magnitudes del espacio y las propiedades físicas de los objetos que lo rodean

7. elige un enfoque determinista o uno aleatorio para el estudio de un proceso o fenómeno y argumenta su pertinencia.

Situación Didáctica 3

Título:	"Atrapado entre las redes "
Contexto:	<p>La OMS ha comenzado a catalogar a las RRSS como un problema de salud pública; ya que según estudios científicos se están arrojando datos alarmantes; por ejemplo, el usar 7 de las 11 redes sociales más populares, está asociado a un riesgo 3 veces mayor de sufrir depresión y ansiedad, en comparación con aquellas personas que solo usan una, dos o ninguna red social. (Fuente: https://prnoticias.com/2019/03/05/rsss).</p> <p>Para hacer inferencia estadística sobre esta situación respecto al comportamiento del estudiantado del COBATAB en cuanto al uso de las redes sociales y poder afirmar o rechazar una hipótesis planteada sobre tal consideración.</p>
Conflicto cognitivo:	<p>d) ¿Influye si el estudiante es de zona urbana o rural en el uso desmedido de las redes sociales?</p> <p>e) ¿Qué relación hay entre el uso de las redes sociales y el promedio de aprovechamiento en el estudiante?</p> <p>f) ¿Crees que como ciudadano estás siendo responsable con el uso que le das a las redes sociales? ¿Por qué?</p>

Propósito de la situación didáctica 3

En equipos de 6 integrantes **construir** una **tabla de contingencia anexa a un reporte con una muestra aleatoria de estudiantes de tamaño $n=100$** que **incluya el número de estudiantes por zona de residencia y la cantidad de las redes sociales que utiliza, según la siguiente clasificación:**

Máximo 2, de 3 a 6, 7 o más

Y utilizando **la prueba Ji cuadrada en un nivel de confianza del 95%** determina "si influye su lugar de residencia para el uso desmedido de las redes sociales", "si tiene efectos sobre su promedio escolar y la influencia que tiene con sus responsabilidades como buen ciudadano". **Para exponerlo ante el grupo y su socialización.**

Nota: se anexa formato de encuesta

Instrumento de Evaluación Situación Didáctica 3

PYE1-B1-B2-GO02 Guía de observación para evaluar la **Situación Didáctica 3**

NOMBRES			Bloque IV Comportamiento de dos variables
			Fecha :

Aprendizaje esperado: Ejemplifica críticamente por medio de modelos matemáticos la relación existente entre dos variables, así como su representación tabular y gráfica, que le permita afirmar o negar hipótesis planteadas en situaciones de su entorno

Situación didáctica 3 "Atrapado entre las redes"	CONTENIDOS ESPECÍFICOS		Grupo		
	<ul style="list-style-type: none"> - Correlación de variables • Regresión lineal • Ji-cuadrada 		<input style="width: 100px; height: 30px;" type="text"/>		
Criterios		%	Cumple		Puntaje
			Si	No	
1. Presenta el producto terminado en el tiempo estipulado por el profesor		10%			
2. Plantea la afirmación o negación de hipótesis en la determinación de la prueba de asociación entre dos variables		15%			
3. Determina la asociación entre las dos variables aplicando la prueba ji cuadrada		30%			
4. Contrasta la afirmación o negación de hipótesis en la determinación de la prueba de asociación entre dos variables		15%			
5. Compara la relación de variables haciendo uso de su representación		20%			
6. Se relaciona con sus compañeros de forma colaborativa mostrando disposición al trabajo.		10%			
TOTAL		100%			
CALIFICACION					

Logros

obtenidos

Aspectos

para

mejorar

Firma del Estudiante

FIRMA DEL FACILITADO



"Educación que genera cambio"
COLEGIO DE BACHILLERES DE TABASCO
PROBABILIDAD Y ESTADISTICA I

Formato encuesta para situación didáctica No.3

"Atrapado entre las redes "

Con la finalidad de conocer la influencia que tiene el lugar de residencia de un estudiante en el uso de las redes sociales y la vez éstas, en su aprovechamiento escolar Contesta las siguientes cuestiones

1. Lugar de residencia

URBANO		RURAL

2. Nivel de escolaridad

SECUNDARIA		BACHILLERATO

3. Promedio de aprovechamiento

6		7		8		9		10

4. Numero de redes sociales que utilizas

Máximo 2		Entre 2 y 6		7 o mas

Nota. *Elaborar en un documento Forms*

Muchas gracias por tu participación

Evaluación Diagnóstica

¿Qué tanto sé? (Apertura)



PYEI-B4-ED03

Evaluación diagnóstica "Análisis de dos variables"

BLOQUE IV: Comportamiento de dos variables

NOMBRE

GRUPO

FECHA

Instrucciones: Lee cuidadosamente cada una de las siguientes cuestiones y elige la opción de respuesta correcta.

1. De las siguientes opciones elige los pares de variables que se relacionan entre si
 - a) El número de hijos y la edad de los integrantes en una familia
 - b) El peso y la estatura de una persona
 - c) La velocidad de un vehículo y el tiempo en recorrer cierta distancia
 - d) Los años de servicio y las inasistencias de una persona en su trabajo
2. Se le conoce así, a la medida que indica el nivel de asociación entre las variables dependiente e independiente
 - a) Coeficiente de determinación
 - b) Coeficiente de correlación
 - c) Diagrama de dispersión
 - d) Recta de regresión
3. A la representación gráfica por medio de puntos en un sistema de ejes coordenados de relación existente de dos variables, se le llama:
 - a) Coeficiente de determinación
 - b) Coeficiente de correlación
 - c) Diagrama de dispersión
 - d) Recta de regresión
4. . Cuando al crecer alguna de las variables, la otra decrece o viceversa, hablamos de una correlación:
 - a) correlación positiva
 - b) correlación negativa
 - c) correlación compleja
 - d) correlación negativa débil
5. Sea $X = 2,4,6,8,10$ Cual será el resultado de la $\sum_{i=1}^3 x_i$
 - a) 30
 - b) 20
 - c) 12
 - d) 6

6. En la ecuación de la recta $y = 3x - 2$ la opción de respuesta que muestra correctamente a la pendiente (m) y la ordenada al origen (b) es

- a) $m=3, b=2$
- b) $m=2, b=3$
- c) $m=3, b=-2$
- d) $m=-2, b=3$

7. Es el análisis que busca probar la asociación o independencia entre dos variables

- a) Correlación lineal
- b) Regresión lineal
- c) Prueba Chi cuadrado
- d) Probabilidad condicional

8. Elige la opción de respuesta que representa una de las causas para lo que se usa la prueba Chi cuadrado

- a) Determinar la dependencia e independencia de dos variables
- b) Determinar la intensidad de la correlación entre dos variables
- c) Determinar la recta de regresión entre dos variables
- d) Determinar la forma de distribución de dos variables

9. En el Chi cuadrado. La _____ contiene los datos obtenidos contados y organizados

- a) Tabla frecuencias absolutas
- b) Tabla de frecuencias relativas
- c) Tabla de frecuencias acumuladas
- d) Tabla de contingencia

10. Es el tipo de variables que intervienen en una prueba chi cuadrada

- a) Aleatorias
- b) Nominales
- c) Discretas
- d) Continuas



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

De la información al conocimiento

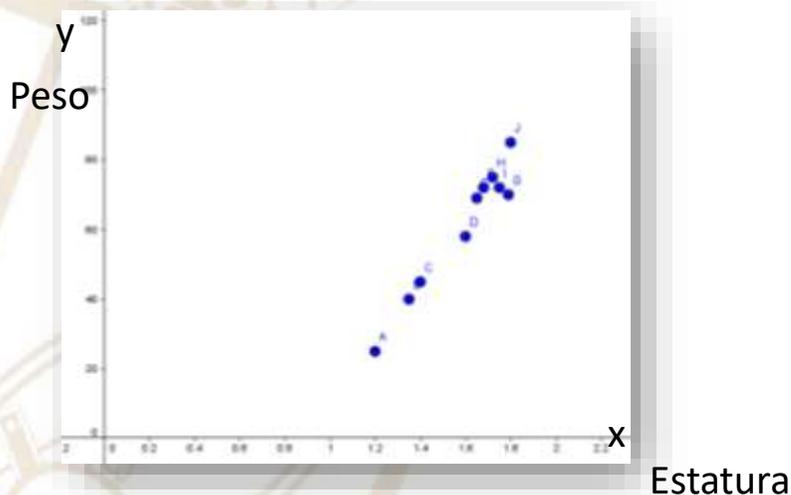


CORRELACION LINEAL

Antes de estudiar el modelo de regresión lineal, a partir del cual podemos analizar la relación entre distintas variables, comenzaremos con su desarrollo usando el estadístico, en particular de dos variables para medir la magnitud de su relación.

Supongamos que el comportamiento del peso de 10 personas de diferentes edades de acuerdo con su estatura se muestra en la siguiente la tabla y grafica.

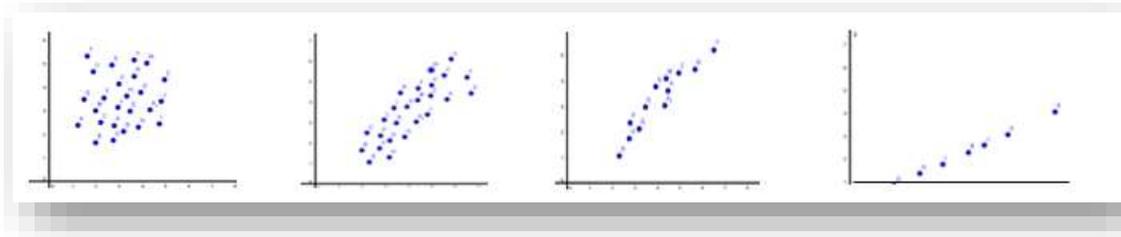
Estatura en mts (x)	1,20	1.35	1.40	1.60	1.65	1.68	1.70	1.72	1.75	1.80
Peso en Kg (y)	25	40	45	58	69	72	70	75	72	85



La distribución de los datos representa a una **tabla de contingencia** y la gráfica que tiene como consecuencia se le denomina **diagrama de dispersión**. Esa grafica es la que muestra si existe o no correlación entre las variables (en este caso peso y estatura de una persona. Dicha correlación puede ser lineal que es nuestro objeto se estudió (o no lineal).

De acuerdo con la tendencia lineal que muestre la gráfica es posible suponer una correlación nula, débil, intensa o perfecta, pudiendo ser positiva o negativa según la variación directo o inverso entre las variables. (las negativas tienden a rectas que bajan de izquierda a derecha

Observa cada tipo de correlación entre dos variables existentes en los siguientes diagramas de dispersión



Nula

Débil

Fuerte

Perfecta

La intensidad de la correlación lineal entre dos variables se mide por su **coeficiente de correlación** y se expresa con la letra “*r*” la fórmula para calcularlo es:

$$\rho_{X,Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

Donde:

- σ_{XY} es la **covarianza** de (X, Y)
- σ_X es la **desviación típica** de la variable X
- σ_Y es la **desviación típica** de la variable Y

De manera análoga podemos calcular este coeficiente sobre un **estadístico muestral**, denotado como r_{xy} a:

$$r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

El valor de r oscila así $-1 \leq r \leq 1$

- $r = 1$ cuando la correlación es positiva perfecta
- $r = -1$ cuando la correlación es negativa perfecta

En ambos casos en el diagrama de dispersión se trata de una serie de puntos que forman una línea recta con pendiente negativa y positiva respectivamente, de modo que:

1. Cuando $r = 0$ las variables no están relacionadas.

PARA APRENDER MÁS:

Coeficiente de correlación de Pearson



https://es.wikipedia.org/wiki/Coeficiente_de_correlaci%C3%B3n_de_Pearson

YOUTUBE:
Video de diagrama de
dispersion_0001



https://es.wikipedia.org/wiki/Coeficiente_de_correlaci%C3%B3n_de_Pearson

2. Cuando r sea cercano a cero por la derecha la correlación entre las variables es débil positiva. Cuando r sea cercano a cero por la izquierda la correlación entre las variables es débil negativa.
3. Cuando r sea cercano a 1 la correlación entre las variables es intensa positiva.
4. Cuando r sea cercano a -1 la correlación entre las variables es intensa negativa.

Para calcular el coeficiente de correlación lineal, que determine el tipo y grado de la intensidad de la relación entre esas dos variables, se recomienda desglosar la expresión del coeficiente de correlación en una tabla como la que se detalla a continuación.

x	y	xy	x^2	y^2
$\sum x =$	$\sum y =$	$\sum xy =$	$\sum x^2 =$	$\sum y^2 =$
$(\sum x)^2 =$	$(\sum y)^2 =$	$n \sum xy =$	$n \sum x^2 =$	$n \sum y^2 =$
$\sum x \sum y =$				

$r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{n s_x s_y}$ $= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$	Operaciones:	Coeficiente de correlación lineal: $r_{xy} =$
--	--------------	--

YOUTUBE:
Coeficiente de correlación
Ejercicio 1



<https://www.youtube.com/watch?v=M9p5VKeVLw>

El objetivo que se persigue cuando se busca la correlación existente entre dos variables asociadas a una situación específica es para poder explicar el comportamiento de una con respecto a la otra y por consiguiente establecer un modelo matemático que mejor se ajuste a la situación, y que nos permita realizar cálculos que nos muestren comportamientos anteriores o a la vez de pronóstico.

YOUTUBE:
Coeficiente de correlación



<https://www.youtube.com/watch?v=7MQbE6VqbB4>

A continuación, encontraras algunas notas importantes y ejemplo de aplicación en contextos de tu entorno, con una mirada hacia la proyección de datos a través de un método consecuente de esta temática Este método se llama **regresión** y para nuestro caso utilizaremos la regresión lineal haciendo uso del método de mínimos cuadrados que se explica en la siguiente sección de lectura.

¿Quieres aprender más?

Da clic al COBACHITO simultáneamente con la tecla CTRL



<https://es.slideshare.net/lexoruiz/regresin-lineal-y-correlacin>

Da clic a COBACHITO simultáneamente
con la tecla CTRL

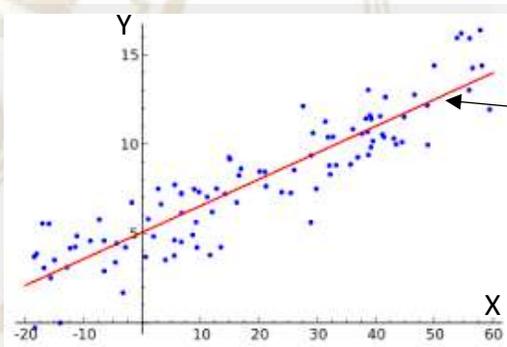


REGRESION LINEAL

Una regresión lineal es una técnica estadística que trata de estudiar la relación entre varias variables estadísticas. Cuando solamente tenemos dos variables diremos que estamos en regresión lineal simple. En investigación, el análisis de regresión utiliza para predecir una de las variables a partir de la otra u otras. Por ejemplo, predecir el éxito que tendrá un estudiante en el examen de admisión para ingresar a la universidad, tomando como base su desempeño en la educación preparatoria, predecir la distancia de frenado de un automóvil que va a cierta velocidad y tiene otras aplicaciones como la calibración de los equipos, la forma que influye la estura del padre sobre la del hijo, la influencia que tiene la presión arterial sobre el ritmo cardiaco de las personas .la proyección del precio de un artículo conforme el paso de los años, etc.

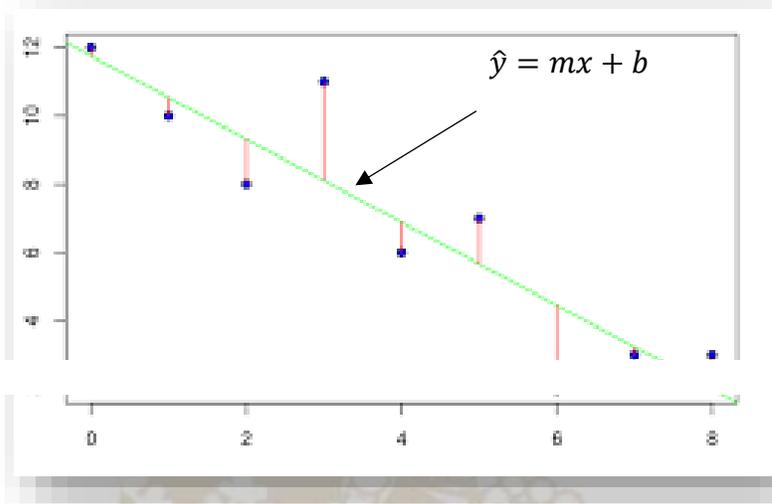
Para ello es muy útil la información que se tenga en el diagrama de dispersión sobre la correlación de las variables en cuestión, ya que da una idea de su comportamiento anterior o posterior mostrado.

Cuando la nube de puntos de un diagrama de dispersión nos informe de una posible correlación lineal, el análisis de regresión lineal tendrá como gran objetivo la predicción de valores para la variable dependiente (Y) a partir de los valores de la variable independiente (X) utilizando para ello una función (una recta de regresión o recta de mejor ajuste) que aproximará lo mejor posible a la nube de puntos.



Recta de
regresión o
mejor ajuste

Para obtener la ecuación de la recta se emplea el método de los mínimos cuadrados, denominado así, ya que considera que la recta de regresión encontrada es la mejor, en sentido de que: la suma de los cuadrados de las distancias verticales de los distintos puntos a la recta de regresión es mínima.



Donde:
 \hat{y} representa el valor estimado
 m es la pendiente
 b la ordenada

Los valores de las constantes que satisfacen el criterio de mínimos cuadrados se obtienen mediante la siguiente fórmula:

$$m = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

$$b = \frac{\sum_{i=1}^n y_i - m \sum_{i=1}^n x_i}{n}$$

$$\hat{y} = mx + b$$

El método consiste en considerar la función que determinaría la suma de todas las distancias verticales (coordenada y_i), elevadas al cuadrado para evitar que las positivas y negativas se contrarresten, entre cada punto y su proyección vertical sobre la hipotética recta. A esta función posteriormente se le calcula dónde alcanzaría el mínimo.

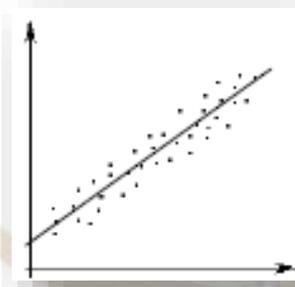
Una vez que sean conocidos los valores de m y b del modelo de regresión lineal simple, éste puede ser utilizado como **modelo predictivo**, esto es, para realizar predicciones de los valores que tomará la variable de respuesta para determinados valores de la variable explicativa. Basta para ello con sustituir en la ecuación de regresión el valor concreto de X que se quiera (X_i). Al hacerlo, se obtendrá el valor predicho para Y según la ecuación de regresión para aquellos casos que en la variable X tomen el valor X_i . Este valor es conocido de forma genérica como puntuación predicha, siendo representado simbólicamente como Y o \hat{y} .

NOTA: La recta de regresión pasa siempre por el centro de gravedad de la nube de puntos, es decir por el punto (\bar{x}, \bar{y})

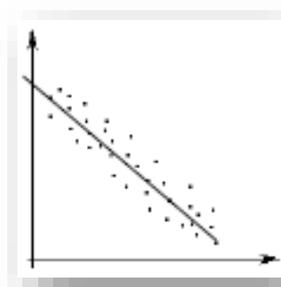
Valores que puede tomar m , es decir nos da información sobre el comportamiento de la variable Y frente a la variable X , de manera que:

- Si $m = 0$, para cualquier valor de X la variable Y es constante (es decir, no cambia).
- Si $m > 0$, esto nos indica que, al aumentar el valor de X , también aumenta el valor de Y .
- Si $m < 0$, esto nos indica que, al aumentar el valor de X , el valor de Y disminuye.

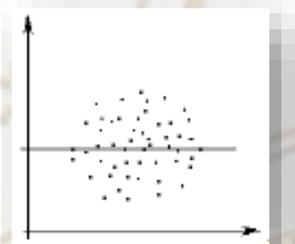
Una interpretación gráfica:



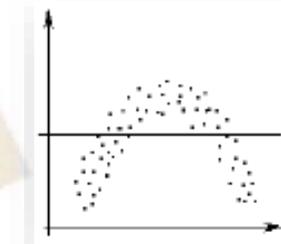
Quando $m > 0$
Relación lineal positiva (directa)



Quando $m < 0$
Relación lineal negativa (inversa)



Ausencia de relación



Relación no lineal: Curvilínea

Para trazar la recta de mejor ajuste en el diagrama de dispersión, es necesario sustituir los valores de X en la ecuación de la recta de regresión lineal, para obtener los correspondiente de \hat{y} . Para ello se requieren dos puntos, los cuales se obtienen eligiendo dos valores de X , cada uno cerca de los extremos del dominio y se calculan los valores correspondientes de \hat{y} .

Puede tomar valores tanto positivos como negativos, siendo mayores en valor absoluto cuanto mayor sea la pendiente de la recta de regresión. Sería igual a 0 si la recta de regresión fuese

horizontal. A continuación, se muestran 4 ejemplos que muestran el vínculo directo entre el valor de m y el tipo de relación existente entre las variables.

Ejemplo:

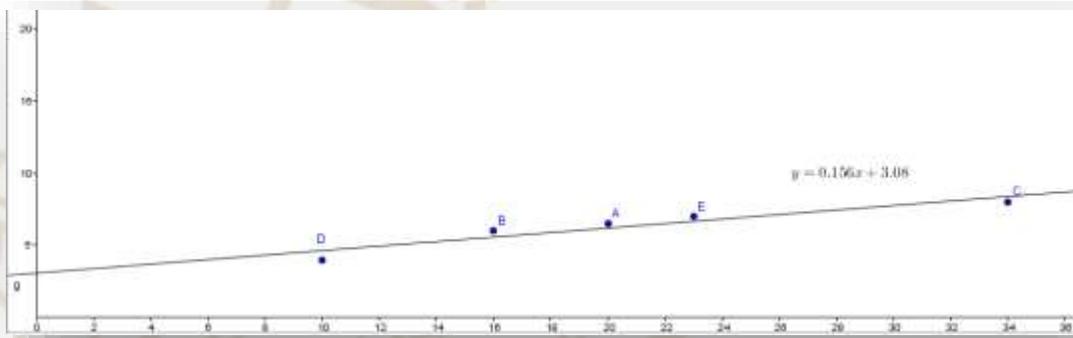
El número de horas dedicadas al estudio de una asignatura y la calificación obtenida en el examen correspondiente de cinco personas es:

Horas de estudio (X)	20	16	34	10	23
Calificación de examen (Y)	6.5	6	8	4	7

- Realiza el diagrama de dispersión
- calcular la recta de regresión de Y sobre X
- Calcular la calificación estimada para una persona que hubiese estudiado 28 horas

Solución:

a)



b) Nos apoyaremos en la siguiente tabla para calcular las

X	Y	XY	X ²	Y ²
20	6.5	130	400	42.25
16	6	96	256	36
34	8	272	1156	64
10	4	40	100	16
23	7	161	529	49



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

$\sum X = 103$	$\sum Y = 31.5$	$\sum XY = 699$	$\sum X^2 = 2441$	$\sum Y^2 = 207.25$
$(\sum X)^2 = 10609$				

Aplicando $m = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$ $m = \frac{(5)(699) - (103)(31.5)}{5(2441) - 10609} = \frac{3495 - 3244.5}{12205 - 10609}$

$$m = 0.156$$

y $b = \frac{\sum_{i=1}^n y_i - m \sum_{i=1}^n x_i}{n}$ $b = \frac{31.5 - (0.156)(103)}{5} = \frac{31.5 - 16.068}{5} = \frac{15.432}{5}$

$$b = 3.08$$

Como la recta de regresiones es $\hat{y} = mx + b$ entonces $y = 0.156X + 3.08$

c) Calificación aproximada si un alumno estudia 28 horas

Aquí $x = 28$

$$\hat{y} = f(x = 28) = 0.156(28) + 3.08 = 7.44$$

$$y = 7.44$$

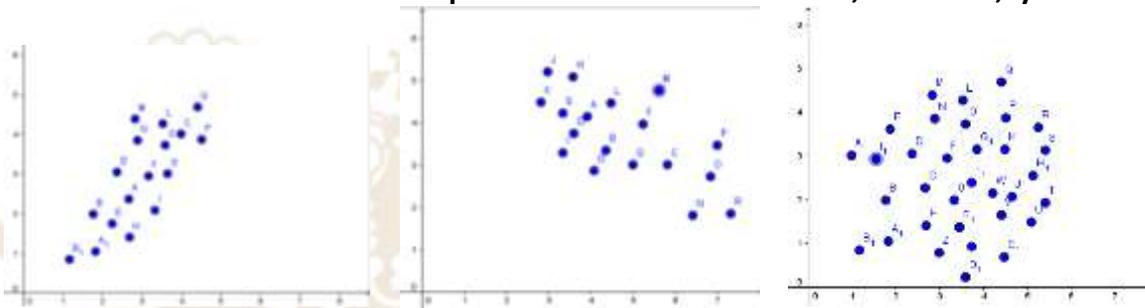
A continuación, se propone como tarea un ejercicio donde se rescatan los conceptos de correlación y regresión lineal y se pide que exploren y visualicen una situación de contexto para hacer una interpretación y estimación de la relación entre las variables involucradas y al a vez una estimación o proyección de una de ellas conociendo la otra.



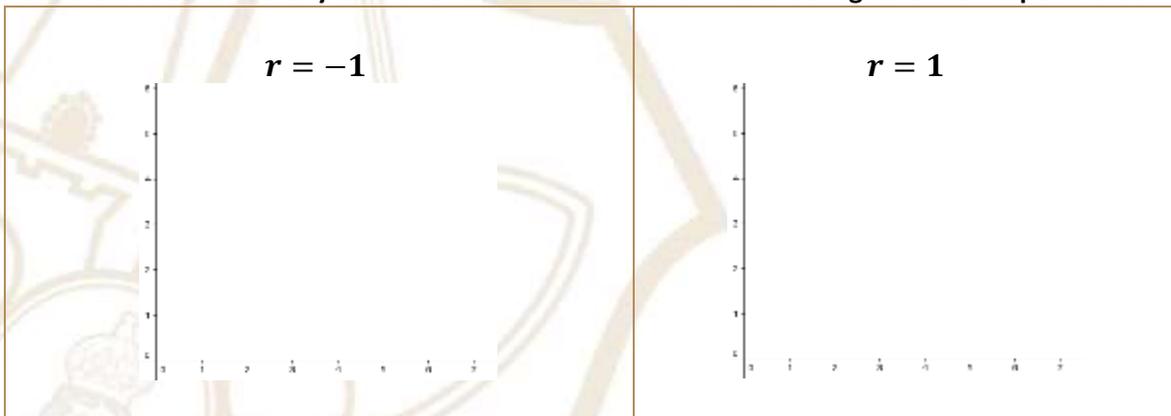
"Educación que genera cambio"

Competencias a desarrollar		TAREA No. 6	Andamio "Correlación y regresión lineal"
CG 5.4 CG 5.5 CG 9.5	CDBM 1 CDBM 2 CDBM 6 CDBM 7		INSTRUCCIONES: Una vez revisadas las lecturas de correlación y regresión, en binas de trabajo completan cada uno de los cuestionamientos establecidos en cada tramo

Observa los siguientes diagramas de dispersión y escribe en cada uno el valor r del coeficiente de correlación correspondiente de entre $0 < r < 1$, $-1 < r < 0$, y $r = 0$



Dibuja los diagramas de dispersión tomando en cuenta la intensidad de la correlación dada entre esas dos variables y escribe en su forma ordinaria a recta de regresión correspondiente



Escribe los modelos matemáticos que usaras para determinar lo que se pide en cada celda

Coeficiente de correlación lineal	Recta de regresión lineal
-----------------------------------	---------------------------



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

Realiza la recolección de la frecuencia o ritmo cardiaco y el peso de 10 familiares o amigos de edades variadas y elabora lo siguiente (analiza quien es la variable dependiente para elegir la Y

<p>Tabla de contingencia</p> <table border="1" style="width: 100%; border-collapse: collapse; margin-top: 20px;"> <tr> <td style="width: 10%; text-align: center;">x</td> <td style="width: 10%;"></td> </tr> <tr> <td style="text-align: center;">y</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>	x										y										<p>Diagrama de dispersión</p>
x																					
y																					

De acuerdo con el diagrama de dispersión elige tu estimación de la intensidad de la correlación entre esas dos variables de las que se muestran a continuación

$0 < r < 1$	$-1 < r < 0$	$r = 0$	$r = 1$	$r = -1$
-------------	--------------	---------	---------	----------

Elige el tipo de correlación entre las variables

Fuerte	Débil	Nula	Negativa perfecta	Positiva perfecta
--------	-------	------	-------------------	-------------------

Utilizando el diagrama de dispersión y determina gráficamente el ritmo cardiaco de una persona que pesa 95kg

Escribe la recta de regresión que obtendrías con la gráfica recordando que es $y = mx + b$ siendo m la pendiente de la recta y b la ordenada al origen .

Para fortalecer tus aprendizajes a cerca de la intensidad de la correlación existente entre dos variables y la forma de hacer una estimación de alguna de ellas conociendo la otra, visita la siguiente dirección electrónica

YOUTUBE:

Ecuación de correlación lineal EJERCICIO
RESUELTO (coeficiente de Pearson, regresión
lineal)



<https://www.youtube.com/watch?v=EE2a2Cr-Jfy>

PYEI-B2-GO03 Guía de observación para evaluar Tarea 06 Ejercicios

NOMBRES		FECHA:

Guía de observación

Aprendizaje esperado: Ejemplifica críticamente por medio de modelos matemáticos la relación existente entre dos variables, así como su representación tabular y gráfica, que le permita afirmar o negar hipótesis planteadas en situaciones de su entorno

TAREITA-6 Andamio Coeficiente de correlación Recta de regresión lineal	BLOQUE IV	CONTENIDOS ESPECÍFICOS - Correlación de variables <ul style="list-style-type: none"> • Representación • Regresión lineal 	Grupo: _____		
Criterios		%	Cumple		Puntaje
			Si	No	
Presenta la resolución del andamio en el tiempo señalado por el profesor		10%			
Determina el valor del coeficiente de correlación a partir de la forma del diagrama de dispersión entre dos variables		30%			
Interpreta la intensidad de la correlación entre dos variables a partir de su diagrama de dispersión		30%			
1. Determina la recta de regresión lineal y realiza proyecciones según el problema de contexto a que se refiere.		20%			
2. Se relaciona con sus compañeros de forma colaborativa mostrando disposición al trabajo.		10%			
TOTAL		100%			
CALIFICACION					

Logros obtenidos

Aspectos para mejorar

Firma del Estudiante

FIRMA DEL FACILITADOR



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

Competencias a desarrollar		TAREA No. 7	Problemario "Correlación y regresión lineal"
CG 5.4 CG 5.5 CG 9.5	CDBM 1 CDBM 2 CDBM 6 CDBM 7		INSTRUCCIONES: Conformar binas de trabajo y apoyados en las lecturas de correlación y regresión, resolver los ejercicios que a continuación se detallan

1. Interprete cada uno de los siguientes coeficientes de correlación y use gráficos de dispersión para representar como se vería cada una de las relaciones entre dos variables (x, y) cualesquiera:

a) $r = -1.0$

b) $r = 0.05$

c) $r = 0.85$

2. La siguiente tabla contiene la edad X y la mínima de la presión sanguínea Y de un grupo de 10 mujeres

Edad	56	42	72	36	63	47	55	49	38	42
Presión	14.8	12.6	15.9	11.8	14.9	13.0	15.1	14.2	11.4	14.1

- Realice el diagrama de dispersión
- Calcule el valor del coeficiente de correlación e interprete el resultado
- Determina la recta de regresión de X sobre Y
- Hacer las siguientes predicciones, solo cuando creas que tengan sentido
 - Presión sanguínea de una mujer de 51 años
 - Presión sanguínea de una niña de 10 años
 - Presión sanguínea de un hombre de 51 años

3. Se realizó un estudio para determinar el efecto que tiene la rapidez de mezclado sobre la cantidad de impurezas en una pintura producida mediante un proceso químico. El estudio arrojó los siguientes resultados.

Rapidez de agitación (rpm)	20	22	24	26	28	30	32	34	36	38	40	42
Cantidad de impurezas (ppm)	8.4	9.5	11.8	10.4	13.3	14.8	13.2	14.7	16.4	16.5	18.9	18.5

4. Estima la cantidad de impurezas aproximadas que se presentan en la pintura cuando se tiene una rapidez de agitación de 33 rpm.

Cobachito te ayuda, visita los sitios que tiene cobachito en su mochila para facilitar la resolución de tu tarea.



Da clic en los logos de YOUTUBE simultáneamente con la tecla CTRL

PYEI-B2-GO04 Guía de observación para evaluar Tarea 07 Problemario

NOMBRES		FECHA

Guía de observación

Asignatura	Probabilidad y estadística I	Comportamiento de dos variables	Grupo	Turno	
Bloque IV		Tarea 07 " Problemario: Correlación y recta de regresión lineal"			
Aprendizajes esperados			Contenidos específicos		
Ejemplifica críticamente por medio de modelos matemáticos, la relación existente entre dos variables, así como su representación tabular y grafica que le permita afirmar o negar hipótesis planteadas en situaciones de su entorno.			<ul style="list-style-type: none"> • Correlación lineal • Regresión lineal 		
Criterios			Cumple		Puntaje
			%	No Si	
1.- Entrega su producto terminado en el tiempo establecido por el facilitador			10%		
2.- Realiza los cálculos adecuados para hallar la recta de regresión lineal			30%		
3.- Representa gráficamente los diagramas de dispersión e identifica la variable dependiente e independiente			30%		
4.- Interpreta las predicciones adecuadamente de los resultados obtenidos			20%		
5.- Se relaciona colaborativamente mostrando disposición al trabajo metódico y organizado			10%		
Calificación					

Logros obtenidos	Aspectos para mejorar
Nombre y firma del Coevaluador	Firma del facilitador



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

De la información al conocimiento



PRUEBA CHI CUADRADO.

En diferentes tipos de eventos donde se da la relación entre dos variables podemos realizar suposiciones del comportamiento de una de ellas conociendo a través de un muestro algunos de los parámetros estadísticos de la otra , y con ello tomar decisiones de aceptar o rechazar esa suposición respecto de una conjetura sobre la dependencia o independencia de esa variable.

Los tipos de variables a las que se hará referencia son del tipo nominal ósea que sus valores representan categorías o grupos en una variable, por ejemplo, para contrastar la preferencia en:

1. Un partido político en una precampaña electoral
2. El uso de cierta red social
3. Color en la compra de un automóvil
4. Gusto por la física ,química y matemáticas
5. Etc.

En todos los casos a través del muestreo y alguna técnica de recolección de información las personas por sexo contestan sus preferencias al respecto, generando lo que se requiere para aceptar o rechazar una conjetura o supuesto

Los supuestos se denominan Hipótesis y el procedimiento para decidir si se aceptan o se rechazan se llama prueba de hipótesis o de significación. Y dentro de la curva de distribución tienen una región específica para su aceptación.

Una prueba de hipótesis es una herramienta de análisis de datos muy importante para la toma de decisiones, que puede en general formar parte de un experimento comparativo más completo entre un supuesto y la realidad.

Los procedimientos que conducen a la aceptación o al rechazo de hipótesis estadísticas como éstas comprenden un área importante de la inferencia estadística.

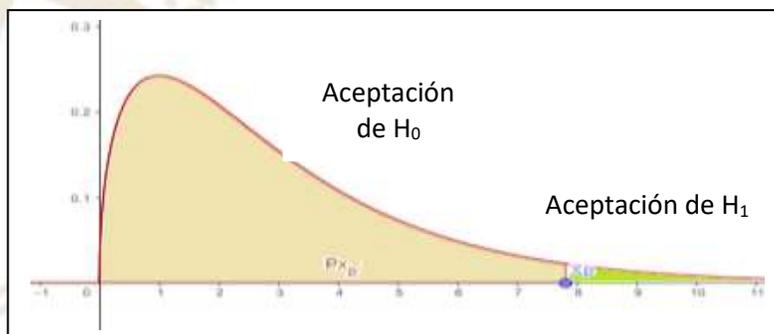
La verdad o falsedad de una hipótesis estadística nunca se sabe con absoluta certeza, a menos que se examine toda la población, lo cual, por supuesto, sería poco práctico en la mayoría de las situaciones. En vez de eso se toma una muestra aleatoria de la población de interés y se utilizan los datos contenidos en ella para proporcionar evidencia que respalde o no la hipótesis. La evidencia de la muestra que es inconsistente con la hipótesis planteada conduce al rechazo de esta.

La estructura de la prueba de hipótesis se establece usando el término hipótesis nula, el cual se refiere a cualquier hipótesis que se desea probar y se denota con H_0 . El rechazo de H_0 conduce a la aceptación de una hipótesis alternativa, que se denota con H_1 .

La hipótesis nula (H_0) indica que ambas variables son independientes, mientras que la hipótesis alternativa (H_1) indica que las variables tienen algún grado de asociación o relación

Para aceptar o rechazar una conjetura o supuesto se aplica el método de prueba y ajuste denominado **CHI CUADRADA**.

La distribución Chi-cuadrada, es una distribución de probabilidad con sesgo positivo



GEOGEBRA:

Comportamiento de la curva en función de los grados de libertad



<https://www.geogebra.org/m/y56cx2>

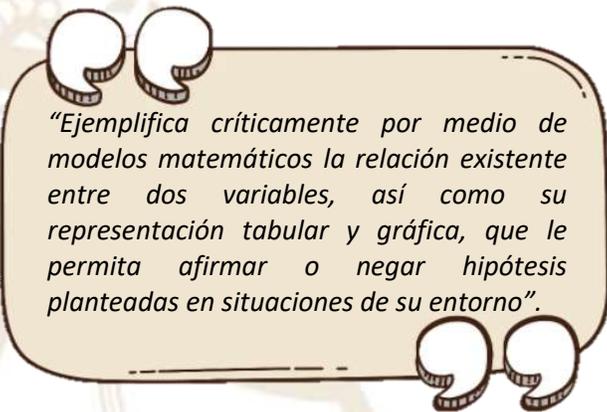
Es una distribución que tiende hacia una curva normal si los grados de libertad de la situación planteada tienden a ser cada vez mayor. Para comprobar el comportamiento de la curva en función de los grados de libertad puedes apoyarte del QR que se encuentra a la izquierda.

La **prueba Chi-cuadrado** (χ^2) o **Ji-cuadrado** fue sugerida por Carl Pearson como una forma de valorar la bondad del ajuste de unos datos a una distribución de probabilidad conocida, y se ha establecido como el procedimiento de elección para el contraste de hipótesis y nos permite probar la supuesta independencia de dos variables cualitativas de una población como una de sus varias aplicaciones en el campo de la inferencia estadística.

Para evaluar la independencia entre las variables, se calculan los valores que indicarían la independencia absoluta, lo que se denomina "frecuencias esperadas", **comparándolos con las frecuencias de la muestra.**

Es un recurso importante, en particular debido a que muchos procedimientos estadísticos en la práctica dependen, en un sentido teórico, de la suposición de que los datos reunidos provienen de un tipo de distribución específico.

Nuestro objeto de estudio dentro de la prueba chi cuadrada es la prueba de independencia entre dos variables cualitativas porque el aprendizaje esperado establecido en nuestro programa de estudio es :



"Ejemplifica críticamente por medio de modelos matemáticos la relación existente entre dos variables, así como su representación tabular y gráfica, que le permita afirmar o negar hipótesis planteadas en situaciones de su entorno".

De ahí que es importante señalar que dentro de los fundamentos en esa prueba se usa la pregunta

¿ la ocurrencia del evento X es independiente a la ocurrencia del evento Y? Con ello puede plantearse las hipótesis nula e hipótesis alternativa de dicha conjetura de acuerdo con cómo se dijo anteriormente lo que estas indican, estableciendo como norma a :

H_0 : La ocurrencia del evento x es independiente del evento y

H_1 : La ocurrencia del evento x no es independiente del evento y

Entendiendo que la independencia de eventos en la teoría de probabilidad se refiere a que la ocurrencia de un evento x no influye en la ocurrencia del evento y , por lo tanto, la probabilidad de que suceda uno no interfiere en la probabilidad de que suceda el otro.

Como a estas hipótesis se les evalúa comparando las frecuencias observadas en la muestra de datos con las esperadas, para ello se construye una tabla de contingencia que permita determinar el valor crítico del chi cuadrado (χ^2) , aplicando el modelo Chi cuadrada bajo cierto nivel de significancia, o margen de error, y los permitidos grados de libertad de la prueba ; parámetros que se explican durante el procedimiento para realizar la prueba que se describe a continuación .

PASOS PARA LA REALIZACIÓN DE UNA PRUEBA DE INDEPENDENCIA CHI CUADRADA

- Realizar una conjetura
- Establecer la hipótesis nula y la hipótesis alternativa
- Calcular el valor de Chi cuadrada

A) Construir una tabla de contingencia de n columnas por m filas

- Obtener las frecuencias observadas (F.O), proveniente de una encuesta, estudio o experimento
- Aplicar la formula $FE = \frac{(total\ de\ columna)(total\ de\ fila)}{gran\ total}$ para obtener las frecuencias esperadas

B) Determinar el valor de la Chi cuadrada con $X^2 = \sum_1^n \frac{(FO-FE)^2}{FE}$

Determinar el valor critico de la Chi cuadrada

- Establecer el nivel de significancia de la prueba (α)
- Calcular los grados de libertad de la prueba(v) con $v = (No.\ columna - 1)(No.\ filas - 1)$
- Buscar en tabla de distribución(Anexa) el valor critico X^2 que es el valor de intersección entre los Grados de libertad y el nivel de significancia de la prueba

Realizar un comparativo entre los valores calculado y crítico de chi cuadrada o construir el área de aceptación y rechazo de la curva grafica de distribución

Aceptar o rechazar la hipótesis nula haciendo uso de las regiones de aceptación en la gráfica de la distribución o considerar que si :

- $x^2_{calculado} < x^2_{critica}$ se acepta H_0
- $x^2_{calculado} > x^2_{critica}$ se rechaza H_0 ,que es lo mismo que aceptar H_1

Para visualizar ejemplos de aplicación de esta prueba en un contexto determinado visita las direcciones electrónicas o escanea el siguiente código y ahí fortalecerás tus aprendizajes

YOUTUBE:

Distribución chi-cuadrado



<https://www.youtube.com/watch?v=gHKMGcn2M5E>

YOUTUBE:

Distribución chi cuadrada
ejemplo resuelto (En 4 pasos)



https://www.youtube.com/watch?v=J9_F_axZlHc

Tabla de Valores críticos de la distribución Chi cuadrada

V	α									
	0.995	0.99	0.98	0.975	0.95	0.90	0.80	0.75	0.70	0.50
1	0.0393	0.0157	0.0628	0.098	0.00393	0.0158	0.0642	0.102	0.148	0.4550
2	0.0100	0.0201	0.0404	0.051	0.103	0.211	0.446	0.575	0.713	1.3860
3	0.072	0.115	0.185	0.216	0.352	0.584	1.005	1.213	1.424	2.3660
4	0.207	0.297	0.429	0.484	0.711	1.064	1.649	1.923	2.195	3.3570
5	0.412	0.554	0.752	0.831	1.145	1.610	2.343	2.675	3.000	4.351
6	0.676	0.872	1.134	1.237	1.635	2.204	3.070	3.455	3.828	5.348
7	0.989	1.239	1.564	1.690	2.167	2.833	3.822	4.255	4.671	6.346
8	1.344	1.647	2.032	2.180	2.700	3.490	4.594	5.071	5.527	7.344
9	1.735	2.088	2.532	2.700	3.325	4.168	5.380	5.899	6.393	8.343
10	2.156	2.558	3.059	3.247	3.940	4.865	6.179	6.737	7.267	9.342
11	2.603	3.053	3.609	3.816	4.575	5.578	6.989	7.584	8.148	10.341
12	3.074	3.571	4.178	4.404	5.226	6.304	7.807	8.438	9.034	11.340
13	3.565	4.107	4.765	5.009	5.892	7.041	8.634	9.299	9.926	12.340
14	4.075	4.660	5.368	5.629	6.571	7.790	9.467	10.165	10.821	13.339
15	4.601	5.229	5.985	6.262	7.261	8.547	10.307	11.037	11.721	14.339
16	5.142	5.812	6.614	6.908	7.962	9.312	11.152	11.912	12.624	15.338
17	5.697	6.408	7.255	7.564	8.672	10.085	12.002	12.792	13.531	16.338
18	6.265	7.015	7.906	8.231	9.390	10.865	12.857	13.675	14.440	17.338
19	6.844	7.633	8.567	8.907	10.117	11.651	13.716	14.562	15.352	18.338
20	7.434	8.260	9.237	9.591	10.851	12.443	14.578	15.452	16.266	19.337
21	8.034	8.897	9.915	10.283	11.591	13.240	15.445	16.344	17.182	20.337
22	8.643	9.542	10.600	10.982	12.338	14.041	16.314	17.240	18.101	21.337
23	9.260	10.196	11.293	11.689	13.091	14.848	17.187	18.137	19.021	22.337
24	9.886	10.856	11.992	12.401	13.848	15.659	18.062	19.037	19.943	23.337
25	10.520	11.524	12.697	13.120	14.611	16.473	18.940	19.939	20.867	24.337
26	11.160	12.198	13.409	13.844	15.379	17.292	19.820	20.843	21.792	25.336
27	11.808	12.878	14.125	14.573	16.151	18.114	20.703	21.749	22.719	26.336
28	12.461	13.565	14.847	15.308	16.928	18.939	21.588	22.657	23.647	27.336
29	13.121	14.256	15.574	16.047	17.708	19.768	22.475	23.567	24.577	28.336
30	13.787	14.953	16.306	16.791	18.493	20.599	23.364	24.478	25.508	29.336
40	20.707	22.164	23.838	24.433	26.509	29.051	32.345	33.660	34.872	39.335
50	27.991	29.707	31.664	32.357	34.764	37.689	41.449	42.942	44.313	49.335
60	35.534	37.485	39.699	40.482	43.188	46.459	50.641	52.294	53.809	59.335

Tabla de Valores críticos de la distribución Chi cuadrada

V	α									
	0.3	0.25	0.2	0.1	0.05	0.025	0.02	0.02	0.005	0.001
1	1.074	1.323	1.642	2.706	3.841	5.024	5.412	6.635	7.879	10.827
2	2.408	2.773	3.219	4.605	5.991	7.378	7.824	9.210	10.597	13.815
3	3.665	4.108	4.642	6.251	7.815	9.348	9.837	11.345	12.838	16.266
4	4.878	5.385	5.989	7.779	9.488	11.143	11.668	13.277	14.86	18.466
5	6.064	6.626	7.289	9.236	11.07	12.832	13.388	15.086	16.75	20.515
6	7.231	7.841	8.558	10.645	12.592	14.449	15.033	16.812	18.548	22.457
7	8.383	9.037	9.803	12.017	14.449	15.033	16.812	18.548	20.278	24.321
8	9.524	10.219	11.030	13.362	15.507	17.535	18.168	20.090	21.955	26.124
9	10.656	11.389	12.242	14.684	16.919	19.023	19.679	21.66	23.589	27.877
10	11.781	12.549	13.442	15.987	18.307	20.483	21.161	23.209	25.188	29.588
11	12.899	13.701	14.631	17.275	19.675	21.92	22.618	24.725	26.757	31.264
12	14.011	14.845	15.812	18.549	21.026	23.337	24.054	26.217	28.3	32.909
13	15.119	15.984	16.985	19.812	22.362	24.736	25.471	27.688	29.819	34.527
14	16.222	17.117	18.151	21.064	23.685	26.119	26.873	29.141	31.319	36.124
15	17.322	18.245	19.311	22.307	24.996	27.488	28.259	30.578	32.801	37.698
16	18.418	19.369	20.465	23.542	26.296	28.845	29.633	32.000	34.267	39.252
17	19.511	20.489	21.615	24.769	27.587	30.191	30.995	33.409	35.718	40.791
18	20.601	21.605	22.76	25.989	28.869	31.526	32.346	34.805	37.156	42.312
19	21.689	22.718	23.900	27.204	30.144	32.852	33.687	36.191	38.582	43.819
20	22.775	23.828	25.038	28.412	31.410	34.170	35.020	37.566	39.997	45.315
21	23.858	24.935	26.171	29.615	32.671	35.679	36.343	38.932	41.401	46.796
22	24.939	26.039	27.301	30.813	33.924	36.781	37.659	40.289	42.796	48.268
23	26.018	27.141	28.429	32.007	35.172	38.076	38.968	41.638	44.181	49.728
24	27.096	28.241	29.553	33.196	36.415	39.364	40.270	42.980	45.558	51.179
25	28.172	29.339	30.675	34.382	37.652	40.646	41.566	44.314	46.928	52.619
26	29.246	30.435	31.795	35.563	38.885	42.923	42.856	45.642	48.29	54.051
27	30.319	31.528	32.912	36.741	40.113	43.195	44.14	46.963	49.645	55.475
28	31.391	32.620	34.027	37.916	41.337	44.461	45.419	48.278	50.994	56.892
29	32.461	33.711	35.139	39.087	42.557	45.722	46.693	49.588	52.335	58.301
30	33.53	34.800	36.25	40.256	43.773	46.979	47.962	50.892	53.672	59.702
40	44.165	45.616	47.269	51.805	55.758	59.342	60.436	63.691	66.766	73.403
50	54.723	56.334	58.164	63.167	67.505	71.420	72.613	76.154	79.490	86.66
60	65.226	66.981	68.972	74.397	79.082	83.298	84.58	88.379	91.952	99.608

Recuperada en junio 9 de 2021 en las páginas 739 y 740 de https://verenciafunez94hotmail.files.wordpress.com/2014/08/8va-probabilidad-y-estadistica-para-ingenier-walpole_8.pdf





TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

Competencias a desarrollar		TAREA No. 8	Problemario "Prueba Chi cuadrado"
CG 5.4 CG 5.5 CG 9.5	CDBM 1 CDBM 2 CDBM 6 CDBM 7		INSTRUCCIONES: Conformados en binas, resuelve los ejercicios de acuerdo con los avances de la lectura PRUEBA CHI CUADRADO.

1. El gerente de ventas de una distribuidora de autos desea conocer si el color de una unidad nueva adquirida por los clientes depende de su género, si se tiene que las ventas logradas durante el primer semestre del año 2020 fue la que se muestra en la siguiente tabla .

GENERO	COLOR DE LA UNIDAD VENDIDA			
	BLANCO	GRIS PLATA	NEGRO	OTRO
MASCULINO	15	25	40	16
FEMENINO	25	20	25	12

Con un nivel de confianza del 95% que conclusión puede tomar el gerente si se sabe que la mayor clientela es de tipo masculino.

2. Durante las presentes campañas electorales una encuesta a aplicada a 1000 personas de la zona rural y urbana en cierto estado de la Republica , muestra las preferencias hacia los diferentes partidos y el partido en el poder de ese estado. Con un nivel de significancia del 5% determina si influye la zona de residencia en la victoria política de un partido en las próximas elecciones . Los datos obtenidos se presentan en la siguiente tabla de contingencia:

ZONA DE RESIDENCIA	PREFERENCIA ELECTORAL	
	PARTIDO EN EL PODER	OTROS
RURAL	450	250
URBANA	200	100

3. Recolecta información sobre el género y color de ojos de tus compañeros de grupo y trata de establecer si hay relación entre el color de los ojos de las personas y su género. Clasifica el color de ojos en negros, cafés y claros. Usa la información obtenida para construir la tabla de contingencia y realiza la prueba con $\alpha = 0.05$

GENERO	COLOR DE OJOS		
	NEGROS	CAFÉS	CLAROS
MASCULINO			
FEMENINO			



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

PYEI-B2-GO04 Guía de observación para evaluar Tarea 08 Problemario

NOMBRES		FECHA

Guía de observación

Aprendizaje esperado: Ejemplifica críticamente por medio de modelos matemáticos la relación existente entre dos variables, así como su representación tabular y gráfica, que le permita afirmar o negar hipótesis planteadas en situaciones de su entorno

TAREITA-08 Problema de contexto	BLOQUE IV	CONTENIDOS ESPECÍFICOS - Chi-cuadrada	Grupo	
Criterios	%	Cumple		Puntaje
		Si	No	
7. Presenta el problema de contexto resuelto en el tiempo estipulado por el profesor	15%			
8. Plantea la afirmación o negación de hipótesis en la determinación de la prueba de asociación entre dos variables	15%			
9. Determina la asociación entre las dos variables aplicando la prueba ji cuadrada	40%			
10. Acepta o rechaza la afirmación o negación de hipótesis en la determinación de la prueba de asociación entre dos variables	15%			
11. Se relaciona con sus compañeros de forma colaborativa mostrando disposición al trabajo.	15%			
TOTAL		100%		
CALIFICACION				

Logros obtenidos

Aspectos a mejorar

Firma del Estudiante	FIRMA DEL FACILITADOR
----------------------	-----------------------



TABASCO

"Educación que genera cambio"



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

PYEI4-B4-PP03		CUESTIONARIO TIPO PLANEA BLOQUE IV
CG. 5.4	CDBM. 1	Indicaciones Lee cuidadosamente cada reactivo e individualmente elige la opción de respuesta correcta
CG. 5.5	CDBM. 2	
CG. 9.5	CDBM. 6	
	CDBM. 7	

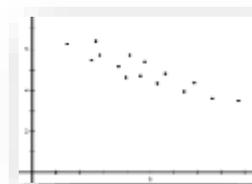
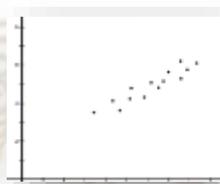
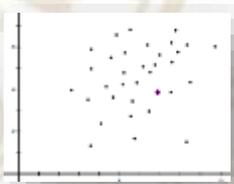
1.- Este tipo de correlación funciona de manera lineal si una de sus variables aumenta, la otra aumenta de la misma manera:

- a) correlación positiva
- b) correlación negativa
- c) correlación compleja
- d) correlación negativa débil

2.- Cuando al crecer alguna de las variables, la otra decrece o viceversa, hablamos de una correlación:

- a) correlación positiva
- b) correlación negativa
- c) correlación compleja
- d) correlación negativa débil

3.- Observa los siguientes diagramas de dispersión y elige la opción que sea correcta, en cuanto al orden y nombre del tipo de diagrama de correlación



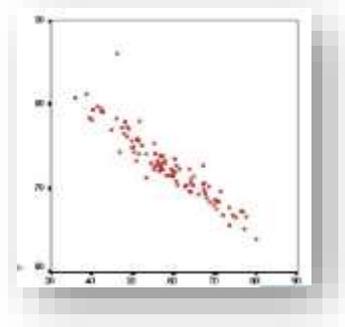
+

- a) 1- Correlación positiva , 2.- Correlación negativa, 3.-Correlación nula
- b) 1.- Correlación negativa, 2.- Correlación positiva, 3.- Correlación nula
- c) 1.- Correlación nula, 2.- Correlación positiva, 3.- Correlación negativa
- d) 1.- Correlación nula 2.- Correlación negativa 3.- Correlación positiva

4. El coeficiente de correlación lineal correspondiente a las variables X e Y cuya distribución conjunta se recoge en el siguiente diagrama de dispersión podría ser:

Fuente: <http://diposit.ub.edu/dspace/html/2445/2921/cuestionari.html>

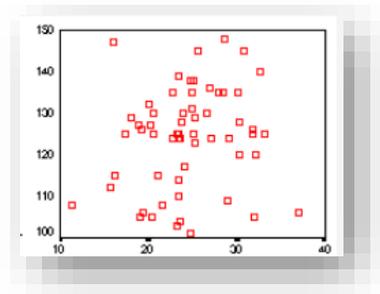
- a) 0.98
- b) - 0.78
- c) 0.23
- d) -1.2



5. El coeficiente de correlación lineal correspondiente a las variables X e Y cuya distribución conjunta se representa en el siguiente diagrama de dispersión podría ser:

Fuente: <http://diposit.ub.edu/dspace/html/2445/2921/cuestionari.html>

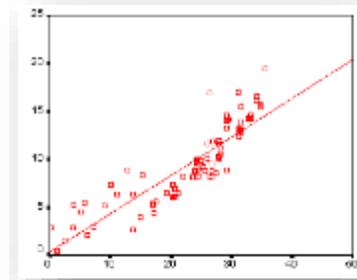
- a) 0.98
- b) - 0.78
- c) 0.23
- d) -1.2



6. La recta de regresión compatible con la siguiente nube de puntos que corresponde a la distribución conjunta de $Y = \text{"Número de respuestas correctas en una prueba"}$ y $X = \text{"Total de horas de asistencia a clase"}$ es

Fuente : <http://diposit.ub.edu/dspace/html/2445/2921/cuestionari.html>

- a) $y = 0.02 - 0.47x$
- b) $y = 0.02 + 0.52x$
- c) $y = 5 + 0.74x$
- d) $y = 5 - 0.04x$



7. Es el comparativo entre la Chi calculado y la Chi critica que permite rechazar la hipótesis nula en una prueba Chi cuadrada

- a) $x^2_{calculado} = x^2_{critica}$
- b) $x^2_{calculado} > x^2_{critica}$
- c) $x^2_{calculado} < x^2_{critica}$
- d) $x^2_{calculado} \neq x^2_{critica}$



8. En el listado anexo se muestra el procedimiento para la realización de una prueba ji cuadrada ,elige la opción que representa la secuencia de pasos para llegar a la toma de decisión (de aceptación o rechazo de una hipótesis),dada una conjetura respecto de alguna situación de estudio.

- a) 6,5,7,1,2,8,4,3
- b) 7,6,1,5,8,4,2,3
- c) 7,6,1,2,5,8,4,3
- d) 6,7,1,2,5,8,4,3

1	Calcular el valor de chi cuadrada (χ^2)
2	Determinar el nivel de significancia y el grado de libertad
3	Aceptar o rechazar la hipótesis nula
4	Interpretar la comparación
5	Obtener el valor crítico de chi cuadrada (χ^2)
6	Escribir la hipótesis nula y la alternativa
7	Realizar la conjetura
8	Realizar una comparación entre el chi cuadrado calculado y el valor crítico

PYEI-B4-MA01 Mapa de aprendizaje para evaluar los Aprendizajes Esperados

Asignatura:	Probabilidad y Estadística I	Bloque: IV	Comportamiento de dos variables	Fecha:	
Nombre				Grupo:	
				Turno:	

Mapa de aprendizaje

1: Necesito ayuda

2: Puedo hacerlo solo

3: Puedo ayudar a otros

Aprendizajes esperados	Nivel			Que debo hacer para mejorar:
	1	2	3	
Ejemplifica críticamente por medio de modelos matemáticos la relación existente entre dos variables, así como su representación tabular y gráfica, que le permita afirmar o negar hipótesis planteadas en situaciones de su entorno				

Estudiante	Facilitador
-------------------	--------------------

Fecha

Referencias BLOQUE IV

Universidad Nacional Autónoma de México Facultad de Estudios Superiores Cuautitlán.

(2021). *PRUEBA DE JI-CUADRADA*.

http://asesorias.cuautitlan2.unam.mx/Laboratoriovirtualdeestadistica/CARPETA%203%20INFERENCIA_ESTADISTICA/DOC_%20INFERENCIA/TEMA%204/08%20PRUEBA%20DE%20CHICUADRADA.pdf

Walpole, Myers, & Myers. (2012). *Probabilidad y estadística para ingeniería y ciencias* (Novena ed.) [Libro electrónico]. Pearson.

https://verenciafunez94hotmail.files.wordpress.com/2014/08/8va-probabilidad-y-estadistica-para-ingenier-walpole_8.pdf

HIMNO COLEGIO

*Oh Colegio de bachilleres impetuosa y querida
institución casa fiel del conocimiento
hoy te canto este himno con amor,
eres rayo de esperanza del mañana
eres la voz de la verdad*

*Oh colegio de bachilleres eres
luz en medio de la oscuridad
//Colegio de bachilleres conducta
clara y firme decisión
Colegio de bachilleres tu misión
para siempre es ser mejor//*

*En Tabasco se ha sembrado
la semilla que un día germinara,
el impulso de la vida modernista
en progreso de toda la sociedad,
es tu memorable historia gran
orgullo para toda la región,
educación que genera
cambio ejemplo digno
en cada generación.*

PORRA INSTITUCIONAL

¡Somos!
¡Somos!

Jóvenes Bachilleres
Jóvenes Bachilleres

Con Valor y Lealtad

De Norte a Sur
De Este a Oeste

Somos Líderes Bachilleres del Sureste
Cobatab Unido, Cobatab Fortalecido

Este encuentro lo gano por que lo gano
Como dijo el peje me canso ganso

¡Somos!
¡Somos!

Jóvenes Bachilleres
Jóvenes Bachilleres

¡Somos!
¡Somos!

Jóvenes Bachilleres
Jóvenes Bachilleres

Cobatab Unido, Cobatab Fortalecido



TABASCO



COBATAB
COLEGIO DE BACHILLERES
DE TABASCO

"Educación que genera cambio"

COBACHITO

Colegio De Bachilleres,
Está de fiesta señores
Pues todos sus estudiantes
Hoy celebran con honores

Que ya llegó la alegría
Es hora de motivar
Bailemos con algarabía
Cobachito nos guiará.

Allá por el acahual
En los rios de tabasco
Aconchado en unas ramas
O nadando sin parar

Un manatí se ha ganado
El cariño de la gente
Cobachito le han llamado
Y no para de bailar.

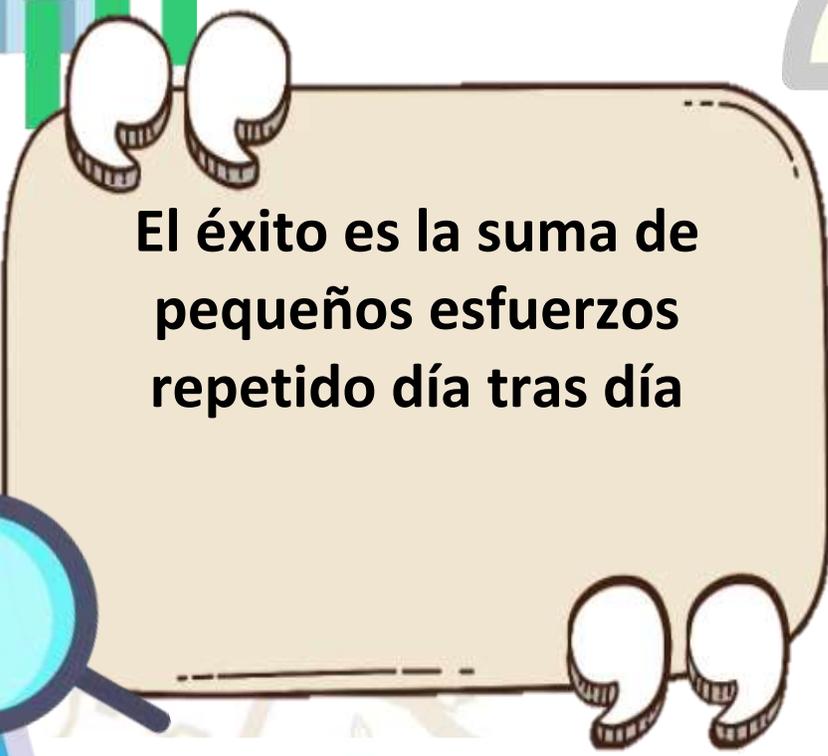
Cobachito, con él vamos a ganar
Cobachito, eres espectacular
Cobachito, respetamos tu hábitat
Cobachito, mascota del cobatab.

Mientras la orquesta se escucha
Y la porra se emociona
Los jovenes bachilleres
A una voz ovacionan.

Con orgullo representan
A una gran institución
Cobatab esta presente
Y cobachito ya llegó.

Cobachito,...





**El éxito es la suma de
pequeños esfuerzos
repetido día tras día**



¡GRACIAS!

Nos vemos en
sexto semestre.



Probabilidad y Estadística I