0032/ 79

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

MATEMÁTICAS DE UN PROYECTO DE INVERSIÓN

T E S I S

QUE PARA OBTENER EL TÍTULO DE

ACTUARIA

P R E S E N T A

MA. TERESA RAMÍREZ ROBLEDO

DIRECTOR DE TESIS: ACT. MARÍA AURORA VALDÉS MICHELL

MÉXICO, D.F. 2003

PAGINACION DISCONTINUA

VNIVERSIDAD NACIONAL AVENMA DE MEXICO

Autorizo a la Dirección General de Bibliotecas UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.
NOMBRE: María Teresa Ramírez Robledo
FECHA: 16 de Enero del 2003
FIRMA: Ma Teresa R. Robledo.

DRA. MARÍA DE LOURDES ESTEVA PERALTA
Jefa de la División de Estudios Profesionales de la Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

Matemáticas de un proyecto de inversión.

realizado por María Teresa Ramírez Robledo

con número de cuenta 8518209 - 9 , quien cubrió los créditos de la carrera de: Actuaria

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis Propietario	Act. María Aurora Valdés Michell.	
Propietario	Act. Marina Castillo Garduño.	
Propietario	Act. Yolanda Silvia Calixto García.	
Suplente	Act. Laura Miriam Querol González.	L.M.Q.G
Suplente	Act. Noemi Velazquez Sánchez.	Noemi Velazquez Sanchez

Consejo Departamental de Matemáticas

M. en C. José Antonio Flores Díaz

FACULTAD DE CIENCIAS
CONSEJO DEPARTAMENTAL
DE
MATEMATICAS

AGRADECIMIENTOS

A MI HERMANA Y MEJOR AMIGA

Que siempre ha estado a mi lado, levantándome si caigo o aplaudiendo mis logros, pero sobre todo, agradezco a dios por haberme dado a mi mejor amiga en mi hermana, quien siempre apoyo este sueño ayudándome a realizarlo.

A LA PROFESORA.

Act. Ma. Aurora Valdés Michell.

Por haberme permitido cerrar un capítulo de mi vida, con la realización de este trabajo, agradeciéndole sobre todo la paciencia, tiempo y dedicación.

A MI TIO

Ing. Juan Robledo Rodríguez

Por haberme guiado, protegido e impulsado donde quiera que estés.

No importa que tan lejos este tu sueño,
Nunca lo abandones.

DEDICATORIAS

A MIS HIJOS.

Gaby y Riky

Quienes son mi inspiración y me brindan la fortaleza que necesito para seguir adelante día a día, y poder ser una buena madre, mujer y maestra.

A MI ESPOSO

Ricardo

Que me ha apoyado durante todos estos años que hemos compartido un hogar.

A MI MADRE

Carolina.

Por ser el pilar de mi vida y el mejor ejemplo ha seguir, pero sobre todo por confiar siempre en mi.

A MIS HERMANOS.

Elizabeth, Sandra y Oscar.

Por todo el apoyo brindado siempre que lo he necesitado.

Y a todos aquellos que me apoyaron en la realización de este trabajo.

CONTENIDO

INTRODUCCIÓN

Actualmente una de las preocupaciones permanentes de las personas, empresas o gobiernos, es tener una mejor calidad de vida la cual se logra mediante la obtención de un cierto nivel de ingresos.

Pero para obtener este cierto nivel de ingresos es necesario " **INVERTIR"**. Lo cual no es nada sencillo y menos para personas que no han tenido experiencia en lo que es un "Proyecto de inversión", decidir ¿en que? y ¿cómo? invertir y el riesgo que implica hacerlo.

El mundo de las inversiones es un mundo interesante. Lo es desde el punto de vista que es de interés para todo el mundo vigilar por la conservación e incremento del patrimonio. Hacerlo es una tarea que requiere, además de contar con los elementos necesarios para ello, esfuerzo, concentración y conocimiento.

Existen diversas áreas en las que se pueden invertir y esto depende de las necesidades y experiencia que se tengan en ellas.

El presente trabajo ha sido realizado con el fin de apoyar matemáticamente en la elección y desarrollo de un proyecto de inversión, específicamente en las áreas de Estadística y Probabilidad obviamente aplicados a un proyecto de inversión.

En el capítulo uno se desarrollan los conceptos básicos, etapas y el riesgo que presenta el elegir un determinado proyecto de inversión.

El uso de las Matemáticas es indispensable en casi todas las áreas del conocimiento humano y en el desarrollo y elección de un proyecto de inversión no es la excepción. Desde la elección y estudio del mercado podemos hacer uso de los conocimientos matemáticos, específicamente de las Estadística, la cual por ejemplo nos ayuda a tomar decisiones sobre el mercado hacia al que iría dirigido nuestro producto y las características de este.

Por otro lado en el capítulo dos introducimos al lector al campo de la Estadística comenzando con las definiciones básicas y posteriormente presentado los métodos numéricos y gráficos utilizados para presentar nuestra información.

En el capítulo tres y cuatro, medidas de tendencia central y medidas de dispersión respectivamente nos ayudan a reforzar las decisiones tomadas y a conocer como se comporta ya sea nuestra muestra o población según sea el caso. Además de explicar claramente el uso de fórmulas de acuerdo a como se presentan nuestros datos , ya sean desordenados o agrupados por variable o intervalos.

Los siguientes capítulos cinco y seis son dedicados a los temas de Probabilidad y Distribuciones de probabilidad temas que nos ayudan por ejemplo a decidir sobre el riesgo que puede presentar un proyecto de inversión con respecto a otro.

Además de que se presentan tres temas importantes del análisis combinatorio que son útiles para resolver problemas de probabilidad que implican una gran cantidad de puntos muestrales.

El último capítulo del presente trabajo tiene como finalidad exponer ejemplos de aplicación de la Estadística y Probabilidad en un proyecto de inversión.

La Probabilidad y la Estadística se relacionan entre sí ya que en esencia la probabilidad es el vehículo que le permite al estadístico usar la información contenida en una muestra para hacer inferencias ó para describir la población de la cual se ha obtenido la muestra.

Los conocimientos de Estadística y Probabilidad, pueden tener aplicación en muchas áreas, desde Demografía hasta Contabilidad o Finanzas, lo importante es tener los conocimientos necesarios en estás áreas para poderlos aplicar de la mejor manera en el momento que sea necesario.

CAPÍTULO I

PROYECTOS DE INVERSIÓN.

1.1 CONCEPTOS GENERALES.

En la actualidad las necesidades económicas que presentan las personas, las empresas o gobiernos, son cada vez mayores por lo cual, los especialistas financieros han desarrollado la capacidad de identificar las oportunidades de inversión que le garanticen la solución a sus problemas económicos, asegurando con ello además la optimización de los recursos con que se cuentan.

Si hablamos de proyectos de inversión, debemos entender primero el significado de cada una de las palabras:

Proyecto: Es un conjunto de datos, cálculos y dibujos articulados en forma metodológica que dan los parámetros de cómo ha de ser y cuanto ha de costar una obra o tarea, siendo sometida a evaluaciones para fundamentar una decisión de aceptación o rechazo.

Inversión : Desde el enfoque económico, se define como el empleo productivo de bienes económicos, que da como resultado una magnitud de éstos mayor que la empleada.
Para el empresario es inversión todo gasto que se efectúa para mantener el funcionamiento o para aumentar el equipo productivo de la empresa.

De las anteriores definiciones podemos desprender la definición de **proyecto de inversión** como: la aplicación de recursos a inversiones fijas que generan ingresos por varios años, es decir es un conjunto de planes detallados que se presentan con el fin de aumentar la productividad de la empresa, para incrementar las utilidades o prestación de servicios, mediante el uso óptimo de fondos en un plazo razonable.

Una definición actual de un proyecto de inversión tiene básicamente tres acepciones:

- **Como documento.** Se refiere a un conjunto de antecedentes relativos a cuatro temas fundamentales que son: aspectos de mercado y comercialización; aspectos técnicos; aspectos jurídicos y administrativos, y aspectos económicos, financieros y sociales, necesarios para tomar una decisión de inversión.

- **Como satisfactor de necesidades.** Esta acepción se refiere a considerar al proyecto como una entidad económica que permite satisfacer una necesidad identificada.

- **Como la parte mínima del presupuesto de capital de una organización.** En este caso, se refiere al concepto empleado en la formulación, análisis y evaluación de cada una de las alternativas de inversión que presenta el presupuesto de capital de una firma.

CAUSAS QUE DAN ORIGEN A UN PROYECTO DE INVERSIÓN.

Los autores que tratan sobre este tema concuerdan en que el origen de los proyectos de inversión es alguno de los siguientes casos o por una combinación de los mismos.

- La existencia de una necesidad insatisfecha.
- La existencia de un recurso susceptible de explotación.
- La existencia de una necesidad política.
- La necesidad de sustituir importaciones.
- La posibilidad de competir a nivel internacional.
- La necesidad de agregar valor a las materias primas.
- La necesidad de mantener la utilidad de productos perecederos.
- La posibilidad de innovar o mejorar productos a menor costo.
- La necesidad de aseguramiento de calidad de un producto.
- La necesidad de responder a cambios en el mercado.

PROBLEMAS, FACTORES Y CONSIDERACIONES.

Se requiere de concebir dos tipos de planes al efectuar la planeación del proyecto que son:

a) **Plan estratégico.-** Deberá comprender las metas (línea de negocio, mercado a satisfacer, etc.), los objetivos generales de la organización que son (económicos y no económicos), las estrategias (decisiones financieras en planeación y control de alto nivel) y las políticas (reglas y principios de aspecto general que son la guía del pensamiento y la actuación)
Los planes estratégicos tienen como finalidad integrar el medio ambiente y la organización como un todo, son orientados a mediano o largo plazo y las estrategias a seguir son decididas por la alta administración.

b) **Plan financiero.-** Consiste en establecer las ideas concebidas mediante cifras que formen los presupuestos, es decir, formular los programas con el objeto de determinar las actividades futuras en tiempo y dinero de una empresa.

Al planear las inversiones, se deberá tomar en cuenta las siguientes consideraciones básicas:

- Aprovechar de manera eficaz, eficiente y redituable los recursos de la organización.
- Buscar soluciones abarcando y satisfaciendo los enfoques técnico y económico ya que son complementarios, para así conseguir los mejores resultados del proyecto en su totalidad.
- Analizar detalladamente las opciones de comprar activos fijos, reparar los existentes o rentar los que sean necesarios según la situación y necesidades de la empresa.
- Ponderar las opciones tecnológicas, detallando sus características y posibles razones para efectuar su selección.

1.2 TIPOS DE PROYECTOS DE INVERSIÓN.

Existen diferentes clasificaciones de proyectos de inversión, la presente tiene como fin adecuarse a los diversos tipos de situaciones que se presenten.

Para efectos de la planeación financiera de las inversiones en bienes de capital, es decir, en inmuebles, maquinaria y equipo, se considera la más razonable la siguiente; siendo esto de acuerdo a:

El tipo de proyecto de inversión:

a) Agropecuarios.- Son todos aquellos dedicados a la producción animal o vegetal.

b) Industriales.- Abarcan la industria manufacturera, extractiva y de transformación relativa a las actividades de la agricultura, pesca y ganadería.

c) De servicios.- Son aquellos que se efectúan para atender necesidades de tipo social como por ejemplo salud, educación, vivienda comunicación etc.

Los resultados a obtener:

a) No rentables.- Son aquellos que no tienen por objetivo obtener utilidades en forma directa.

b) No medibles.- Se denomina en esta forma a aquellos cuyo objetivo es lograr una utilidad en forma directa, siendo difícil cuantificar la misma.

c) De reemplazo.- La finalidad es sustituir activos debido a desgaste u obsolescencia de éstos, logrando así mantener la eficiencia de la planta productiva.

d) De expansión.- Tienen como objetivo lograr una mayor capacidad productiva mediante el reemplazo del equipo por ser obsoleto o por la modernización del mismo para obtener eficiencia y de esta manera poder hacer frente a la tendencia creciente de ventas en una empresa en proceso de desarrollo, o bien porque la empresa desea ganar mayor mercado.

Su naturaleza:

a) Dependientes.- Son aquellos que se encuentran condicionados entres sí, en otras palabras, si se tienen tres proyectos A, B y C, la aprobación de uno de ellos solo será posible si los otros dos también son aceptados.

b) Independientes.- Se denominan así, puesto que la aprobación de uno de ellos no descarta la posibilidad de la aceptación posterior de cualquiera de los restantes, es decir, la aprobación del A no influye en la adquisición de B y C, ya que el objetivo de cada uno de ellos es distinto.

c) Mutuamente excluyentes.- Son aquellos cuya finalidad o función a realizar dentro de la empresa es la misma, por ésta razón la aceptación de uno de ellos provoca la eliminación de los restantes.

Otras personas prefieren clasificar los proyectos en función de la entidad que los promueve y así hablan de proyectos sociales y proyectos privados.
Actualmente el concepto se emplea también dentro del ámbito de la denominada Ingeniería financiera, donde se habla de proyectos de restauración de pasivos, proyectos de mexicanizaciones, proyectos de emisión de deuda, proyectos de capitalización de pasivos etc.

1.3 ETAPAS DEL PROYECTO.

En la elaboración de proyectos de inversión para bienes de capital dependiendo de su complejidad y magnitud se pueden considerar diversas etapas de análisis y evaluación, por lo general, se destacan seis básicas:

1) Estudios preliminares.- Son aquellos que sirven de preámbulo para analizar posteriormente en forma sólida un proyecto, se basan en la información que se tiene a la mano.
Dentro de esta primera etapa se busca conceptuar la idea del proyecto en forma general, tratando de delimitar los rangos máximos y mínimos de la inversión.

Cabe señalar que si la empresa desea crecer y desarrollarse deberá promover e impulsar un ambiente creativo, es decir un medio ambiente en el que existan las condiciones adecuadas para fomentar la iniciativa del personal.

2) Anteproyecto.- Se conoce también como " estudio previo de factibilidad", consiste en comprobar mediante información más detallada (Estadísticas macroeconómicas, existencia de recursos propios, fuentes de financiamiento, incentivos fiscales, magnitud de la competencia, identificación del consumidor potencial mediante pruebas de mercado, etc.), presentando la semblanza del proyecto, rendimiento esperado y un pronostico de los recursos financieros, humanos, y técnicos necesarios a través de un trabajo final.

3)Constitución del comité.-Se debe formar un grupo de trabajo interdisciplinario, esto es, establecer un comité del proyecto, en el cual estarán definidas las tareas y responsabilidades y niveles de autoridad en función del proyecto que se trate.

4) Estudio de factibilidad.- En esta cuarta etapa se realiza el "Documento del proyecto", se encuentra integrado por los análisis de mercado, ingeniería, económico-financiero y el plan de ejecución; es aquí donde se establecen los elementos cuantificables y no cuantificables de un proyecto, además de la combinación adecuada de estos.

5) Puesta en marcha y funcionamiento normal.- Se refiere a la implementación del proyecto, dentro de este contexto se encuentra la compra del bien, su instalación, capacitación del personal, operación, mantenimiento, etc.

6) Control.- Consiste en la comparación y medición de los resultados reales contra los presupuestados (análisis de variaciones), lo cual puede realizarse en forma parcial o total, teniendo como objetivo corregir o mejorar la actuación del proyecto.

1.3.1 DOCUMENTO DEL PROYECTO.

El objetivo del documento del proyecto es proveer los elementos necesarios para la toma de decisiones en lo concerniente al apoyo que se deberá prestar a la ejecución del proyecto de inversión en el bien del capital. El documento del proyecto esta formado por:

a) Análisis de Mercado.
b) Análisis de Ingeniería.
c) Análisis Económico-Financiero.
d) Plan de Ejecución.

1.4 RIESGO EN UN PROYECTO DE INVERSIÓN.

La gestión, empresarial en la toma de decisiones sobre la evaluación de proyectos de por sí compleja, ve acentuada su complejidad en virtud de diversos aspectos y fenómenos que interactúan en su entorno, entre los que destacan: el alza y variabilidad de las tasas de interés en el mercado de capitales, la inflación interna e importada y los problemas sociales y políticos, tanto internos como externos.

Ante esta realidad y su creciente perspectiva, la toma de decisiones en la empresa requiere de respuestas ágiles y certeras que implican el dominio y aplicación de conceptos de diversas disciplinas, entre otras las del tipo: económico, financiero, matemático, estadístico, mercadológico, informático, fiscal y administrativo.

Tales conceptos y conocimientos, a su vez requieren de un enfoque interdisciplinario y de sistemas que permitan optimizar el logro de los objetivos planteados en la toma de decisiones; y es precisamente en este contexto donde el tratamiento y medición del riesgo que involucran los proyectos de inversión cobran una importancia radical que merece una atención especializada.

La definición de riesgo que se encuentra en un diccionario es:

- El peligro en relación con algún daño probable.
- La falta de conocimiento seguro sobre un evento.
- Cada una de las contingencias que pueden ser objeto de un contrato de seguro.

Además de los diversos significados de la palabra riesgo, dicho concepto tiene otras más en función de la disciplina en la cual se emplea, y en función del tema financiero de que se trate.

- **Especialista en finanzas:** El riesgo es la variación de un rendimiento, valor o utilidad esperada.
- **Contador público:** El riesgo son los errores e irregularidades que se pueden cometer al tomar una decisión.
- **Administrador de empresas:** El riesgo está representado por la toma de decisiones en condiciones inciertas.
- **Un programador:** El riesgo se refiere a la toma de decisiones en la que al menos una variable de decisión es arbitraria.

Aunque además de las diversas acepciones de riesgo que pueden dar los profesionales, los textos de finanzas incluyen las siguientes definiciones.

- **Riesgo comercial:** Causado por fluctuaciones de utilidades de operación. Este tipo de riesgos depende de la variabilidad en la demanda, el precio de venta, de la estructura de costos y gastos y del grado de apalancamiento operativo de la organización o proyecto de inversión.

- **Riesgo financiero:** La incertidumbre de los rendimientos futuros para los socios de una firma, misma que deriva de la estructura financiera adoptada por ésta, es decir de la manera en que se haya decidido el financiamiento de los activos de la empresa.

- **Riesgo de mercado:** Se refiere al riesgo que se tiene de que el precio de una acción cambie, debido a las variaciones en el entorno del mercado bursátil en general, o a que los precios de todas las acciones tienen cierto grado de correlación con las fluctuaciones generales del mercado de valores.

- **Riesgo sistemático:** Este riesgo es inevitable, independientemente del grado de diversificación de la cartera de proyectos o inversiones que un inversionista pueda lograr, este deriva del hecho de que hay otros peligros en el conjunto de la economía que amenazan a todos los negocios. Siendo ésta la razón por la que los inversionistas están expuestos a la incertidumbre del mercado, sin importar el número de acciones que posean. Para un inversionista que posea una cartera razonablemente variada, lo más importante será el riesgo de mercado, por ésta razón para un inversionista que diversifica el riesgo la fuente de incertidumbre predominante radica en la variabilidad del mercado, ya que éste arrastrará con él a la cartera de inversionistas.

1.4.1 TÉCNICAS PARA LA MEDICIÓN DEL RIESGO EN UN PROYECTO DE INVERSIÓN.

Para la evaluación del riesgo de un proyecto de inversión o de una cartera o (portafolios) de proyectos, se puede realizar básicamente de tres maneras:

a) Mediante un criterio totalmente informal y subjetivo.
b) Mediante consideraciones formales, las cuales pueden ser, tanto de carácter subjetivo, como objetivo.
c) Mediante una combinación de ambos criterios.

En el caso de criterios informales, sus procedimientos de valuación son intrínsecamente débiles, ya que los factores que afectan el riesgo no son examinados en forma explicita, ni de manera sistemática.
Por ejemplo teniendo dos proyectos de inversión A y B, en donde los porcentajes de rendimientos esperados son de 32% y 18% respectivamente, y asumiéndose que el proyecto A es más arriesgado que el B, aunque no se haya realizado un

estudio para medir el grado preciso de riesgo en cada caso, entonces se está hablando de un criterio informal.

En el caso de criterios formales, la administración financiera emplea diversas técnicas que permiten la medición y evaluación del riesgo, tanto a nivel de alternativas individuales, como para combinaciones de proyectos (carteras o portafolios), tales como: Análisis de sensibilidad, **distribuciones de probabilidad**, modelo de valoración de activos de capital, coeficiente de volatilidad, tasas de descuento ajustadas por el riesgo, **árboles de decisión**, equivalentes de certeza y modelos de simulación.

La determinación del riesgo mediante la técnica de distribuciones de probabilidad, requiere el empleo de algunos conceptos estadísticos tales como: **La media, la desviación estándar, la estandarización de valores y la búsqueda de valores bajo la curva normal.**

Por otro lado los árboles de decisión o probabilidad son gráficos asociados a las **probabilidades de ocurrencia** en diferentes escenarios, para uno o varios periodos, que permiten tener una idea visual sobre un problema de inversión.

Los siguientes capítulos pretenden reforzar la aplicación de conceptos de estadística y probabilidad relacionados con temas básicos como: estudios de mercado, evaluación de mercados en condiciones de riesgo, etc., además de reforzar el estudio de la estadística, sus vínculos y aplicabilidad en las disciplinas y temas de su especialidad o en aquellas que les sean afines.

CAPÍTULO II

ELEMENTOS BÁSICOS DE LA ESTADÍSTICA.

2.1 DESARROLLO HISTÓRICO DE LA ESTADÍSTICA.

Estadística Etimológicamente significa " Ciencia del Estado", pues en un principio su campo de acción era la recopilación y análisis de datos relativos a la población y riqueza que la administración de los Estados exigían para fines de **la Guerra** y de **las Finanzas.**

Actualmente la Estadística se emplea como un valioso instrumento en diversos campos de las ciencias y en las más variadas actividades, de tal manera que es difícil encontrar una fase de la actividad humana que no considere útil cuando menos ocasionalmente la utilización de la Estadística.

2.2 NECESIDAD DE LA ESTADÍSTICA.

La estadística o métodos estadísticos actualmente están desempeñando un importante papel en todas las facetas del proceso humano, como un valioso auxiliar en los más diversos campos del conocimiento y en las más variadas actividades de las ciencias fundamentales y aplicadas. Es una materia básica en las carreras de Sociología, Economía, Administración de Empresas, Pedagogía, Psicología, Ingeniería Civil, Historia Medicina, Física y otras más.

A partir de la mitad de este siglo, la necesidad de utilizar los datos numéricos para expresar la información cuantitativa, ha adquirido gran importancia económica y sociológica. Por ejemplo hace 30 años cualquier persona se identificaba exclusivamente por sus nombres y apellidos. Hoy en día tiene otros datos asociados con el nombre: Número del acta de nacimiento, Número del IMSS etc.
Entre las muchas aplicaciones que tiene la Estadística destacaremos las siguientes:

- Representación funcional de los fenómenos numéricos.
- Representación gráfica de situaciones de toda índole: Económica, poblacional, física, biológica, demográfica, etc.
- Recopilación de datos y su interpretación.
- Predicción científica de efectos de causas evaluables de una forma matemática.
- Interpretación de la experimentación.
- Aprovechamiento del muestreo para el conocimiento de los fenómenos.

2.3 DIVISIÓN DE LA ESTADÍSTICA.

La Estadística es el conjunto de técnicas que se emplean para la recolección, organización, análisis e interpretación de datos. Los datos pueden ser cuantitativos, con valores expresados numéricamente, o cualitativos en cuyo caso se tabulan las características de las observaciones.

La Estadística se divide en:

a) **Estadística descriptiva:** Comprende las técnicas que se emplean para resumir y describir datos numéricos. Estos métodos pueden ser numéricos o gráficos.

   Ejemplo : El promedio mensual de ventas de un circuito digital se puede describir mediante el uso de gráfica de barras. Obteniéndose la media aritmética de las ventas mensuales de este dispositivo.

b) **Estadística inferencial:** Comprende las técnicas con las que con base únicamente en una muestra sometida a observación, se toman decisiones sobre una población. Dado que estas decisiones se toman en condiciones de incertidumbre es necesario el uso de conceptos de probabilidad.

   Ejemplo: Para estimar el voltaje requerido para provocar fallas en un dispositivo eléctrico, una muestra de estos dispositivos puede someterse a voltajes crecientes, hasta que falle cada uno de ellos. Con base en estos resultados muestrales puede estimarse la probabilidad de falla a varios niveles de voltaje de los demás dispositivos de la población muestreada.

A continuación se darán definiciones elementales:

UNIVERSO: Conjunto determinado de objetos o elementos.

POBLACIÓN: Conjunto de todas las mediciones de interés para quien obtiene la muestra y cuentan con características en común. La población es un subconjunto del universo.

MUESTRA: Es un subconjunto de mediciones seleccionadas de una población de interés.

El universo puede constar de un solo objeto, por ejemplo un automóvil. La población de ese universo pueden ser todas las mediciones del espesor de las piezas del automóvil. También pueden ser dichas piezas en cuanto, a las substancias de que están hechas.

Una muestra es aleatoria si al elegirla, cada elemento de la población tiene la misma oportunidad de ser elegido. " Solo con esta clase de muestra es posible aplicar las leyes de la probabilidad matemática".

2.4 VARIABLES ALEATORIAS.

Las observaciones generadas por un experimento caen en una de dos categorías, a saber, cuantitativas o cualitativas por ejemplo la producción diaria de una planta industrial es una observación cuantitativa o numéricamente medible, por otro lado las descripciones del clima, tales como lluviosos, nublado o soleado son cualitativas. En algunos casos es posible convertir los datos cualitativos a cuantitativos, asignando un valor numérico a cada categoría para formar una escala. La producción industrial a menudo se clasifica como de primera clase, segunda clase, etc.
Los sucesos de interés asociados con el experimento son los valores que los datos pueden tomar y son por lo tanto sucesos numéricos.

Una variable aleatoria: Es una función con valores numéricos definida sobre un espacio muestral.

Existen por lo tanto dos tipos de variables aleatorias respecto de los datos obtenidos en el muestreo.

a) **Variable aleatoria discreta:** Esta variable se representa como individualidades. Solo puede asumir un conjunto numerable de valores, si se pueden enumerar entonces se trata de una variable aleatoria discreta. Por ejemplo se tienen 3 ó 4 hijos, no cuatro y medio.

b) **Variable aleatoria continua:** Se miden entre dos valores cualesquiera, en donde hay por lo menos un intermedio, ésta variable puede asumir el número infinitamente grande de valores correspondientes a los puntos sobre un intervalo en una línea recta. Por ejemplo un caballo puede hacer un recorrido diario de 2.6 Km. valor que está entre 2 Km. y 3 Km.

A continuación se dan algunos ejemplos de variables aleatorias:

1.- La altura del agua en una represa **Var. Ale. Contínua.**

2.- El número de persona extraviadas en un mes. **Var. Ale. Discreta.**

3.- Su presión arterial. **Var. Ale. Discreta.**

4.- La cantidad de lluvia que cae en la ciudad de México en una semana **Var. Ale. Contínua.**

5.- La altura de las personas en un salón. **Var. Ale. Contínua.**

6.- la distancia que recorre un alumno de su casa al colegio. **Var. Ale. Contínua.**

2.5 ELEMENTOS ESCENCIALES DE UN PROBLEMA ESTADÍSTÍCO.

El objetivo de la Estadística es hacer inferencias (predicciones, decisiones) acerca de una población, sobre la base de la información contenida en una muestra. Para conseguir el objetivo tendremos que cada problema debe contener los siguientes elementos para poder obtener los mejores resultados en la investigación:

1.- Plan de recopilación: Es el conjunto de procedimientos utilizados en la planeación de la recopilación, es decir, representa un trabajo analítico que tiende a determinar lo que hay que buscar, en donde y como obtenerlo. Es necesario planear una serie de preguntas cuyas contestaciones interesen para que sean contestadas sin dificultad.

2.- Recopilación de los datos: Es el conjunto de operaciones a realizar para obtener los datos de los que nos interesa del fenómeno a estudiar.

3.- Crítica de los datos: De esta etapa depende que los datos se apeguen a la verdad y usar los datos, si están contestadas todas las preguntas y si están acorde con la demás información.

4.- Clasificación y recuento de los datos: Consiste en agruparlos según los caracteres que se manifiesten en el fenómeno.

5.- Representación de los datos: Consiste en mostrar los datos obtenidos, debidamente ordenados en forma clara y precisa con el objeto de facilitar la interpretación y análisis de los mismos. La presentación puede hacerse en forma escrita, en tablas, cuadros estadísticos o gráficas.

6.- Análisis de los datos: Consiste en detectar las diferentes variaciones para poder tomar decisiones finalmente.

2.6 ESTADÍSTICA DESCRIPTIVA.

Una vez que se han presentado las principales definiciones relacionadas con la Estadística descriptiva comenzaremos por presentar la toma de los datos obtenidos en la encuesta.

2.6.1 TOMA DE DATOS Y ORDENACIÓN.

Toma de datos: Es la obtención de una colección de información que no ha sido ordenada numéricamente.

Ejemplo: La cantidad de cortos circuitos que presenta diariamente un tablero electrónico sometido a altos voltajes (Tomados durante 10 días)

Día	1	2	3	4	5	6	7	8	9	10
Cortos	0	1	3	2	0	4	1	1	0	3

Se observa que al realizar nuestro muestreo los datos que se obtienen de él se encuentran en desorden, lo cual no facilita el estudio del fenómeno, esto exige que se presenten de otra manera los datos obtenidos por lo cual recurrimos a la ordenación de la variable.

Ordenación: La ordenación de la variable estadística puede ser en forma ascendente o descendente; y es la colocación de los mismos en forma ordenada.

Ejemplo: Tomamos como población al conjunto de valores de la edad de los alumnos que logran egresar como técnicos electricistas. En donde la muestra es la siguiente:

19	21	23	19	20	25
25	22	22	27	26	31
31	33	31	30	24	24
21	20	20	19	25	23
27	28	29	33	20	20

De la forma en que están presentados los datos, no es posible establecer ninguna relación útil, organizándolos en una tabla nos darán una mejor información. Por lo regular, siempre se ordenan en forma ascendente.

Ordenados en un cuadro quedarán así:

Observando la tabla, ahora es mucho más fácil interpretar los datos obtenidos.

De los 30 alumnos egresados tenemos los siguientes datos:

5 alumnos tienen 20 años.
3 alumnos tienen 19 años.
3 alumnos tienen 25 años.
2 alumnos tienen 27 años.
1 alumno tiene 30 años.

Estos datos fueron tomados de alumnos que terminaron sin haber reprobado ninguna materia.

19	19,19,19	3
20	20,20,20,20,20	5
21	21,21	2
22	22,22	2
23	23,23	2
24	24,24	2
25	25,25,25	3
26	26	1
27	27,27	2
28	28	1
29	29	1
30	30	1
31	31,31,31	3
33	33,33	2
		30

Sin embargo ésta forma de presentar los datos es inaplicable para cuando se cuenta con una gran cantidad de información que hay que analizar. Una forma más sencilla es por medio del uso de una tabla de frecuencias.

2.6.2 DISTRIBUCIÓN DE FRECUENCIAS.

Antes de comenzar la construcción de una tabla de distribución de frecuencias se darán las definiciones asociadas a este tema.

Distribución de frecuencias ó Tabla de frecuencias: Es la ordenación tabular de los datos en clases y las frecuencias correspondientes a cada una.

Valor mayor: Entre los datos obtenidos, o de los datos de la muestra que se haya tomado al azar, determinar el que valga más.

Valor menor: Dato menor de la muestra o de la recopilación.

Rango o Recorrido: Es la diferencia entre los dos valores anteriores.

$$R = V_{mayor} - V_{menor}$$

Frecuencia absoluta: La cantidad de veces que aparece o se repite un dato.

Frecuencia relativa: La cantidad de casos o elementos de una clase o intervalo dividida entre el total de frecuencias.

Intervalo: División generalmente igual del rango, cada uno se llama clase.

Límite inferior del intervalo: Valor inicial del intervalo.

Limite superior del intervalo: Valor final del intervalo.

Marca de clase: Promedio aritmético de los límites del intervalo.

Aplicaremos las definiciones vistas a un ejemplo.

Ejemplo: 15 Técnicos electricistas interesados en obtener una beca para un curso de mantenimiento presentaron un examen diseñado para medir su aptitud. (El puntaje máximo es de 100 aciertos), siendo los resultados los siguientes:

79	92	86	76	92
88	86	76	63	46
87	88	87	92	79

Los encabezados de la tabla serán: Variable (Es el puntaje obtenido en el examen), frecuencia y frecuencia relativa. A esta manera de representar los datos se le conoce como **distribución de frecuencias.**

VARIABLE	FRECUENCIA	FRECUENCIA RELATIVA
46	1	1/15= .066
63	1	1/15= .066
76	2	2/15=.133
79	2	2/15=.133
86	2	2/15=.133
87	2	2/15=.133
88	2	2/15=.133
92	3	3/15=.200
TOTALES	**15**	**.998**

Se observa en la tabla que el total de frecuencias coincide con el número de datos, así como el total de la frecuencia relativa siempre debe de ser 1 ó .999 de lo contrario alguna operación ha sido mal calculada.

La tabla anterior puede tener otra presentación **acumulando** las frecuencias, tanto la absoluta como la relativa, siendo los encabezados de la tabla: Variable, Frecuencias acumuladas y Frecuencia relativa acumulada (Estas columnas se obtendrán realizando la suma de las frecuencias anteriores y la que ya se tenia).

Por ejemplo: El renglón uno se mantiene igual, el renglón dos se obtiene de la suma dela frecuencia anterior y la frecuencia de ese renglón 1+1 =2, para el tercer renglón se tendrá 2+2 =4 etc. Realizando lo mismo pero para la frecuencia relativa.

VARIABLE	FRECUENCIA ACUMULADA	FRECUENCIA RELATIVA ACUMULADA
46	1	.066
63	2	.133
76	4	.266
79	6	.399
86	8	.532
87	10	.665
88	12	.798
92	15	.998
TOTALES	15	998

Las tablas anteriores aunque presentan de una mejor manera los datos se pueden condensar distribuyendo las frecuencias por medio del uso de intervalos.

2.6.3 INTERVALOS Y LÍMITES REALES DE CLASE.

Para formar una **Distribución de frecuencias por intervalos** (llamados clase) deben de ser del tipo $a \angle x \leq b$ o sea (a, B]

El tamaño t del intervalo lo obtendremos de la siguiente manera:

- Primero se obtiene el rango de la muestra.
  Rango = V_{mayor}-V_{menor}
- Se divide el rango entre el número de intervalos deseados

  t= Rango $\div$Num. de intervalos deseados.

Tomando el ejemplo anterior, construiremos una distribución de frecuencias con 6 intervalos:

Rango = 92-46 = 46

t= 46/6 =8.3 que por conveniencia lo hacemos 8, es decir, t = 8

La tabla de intervalos de clase queda de la siguiente forma:

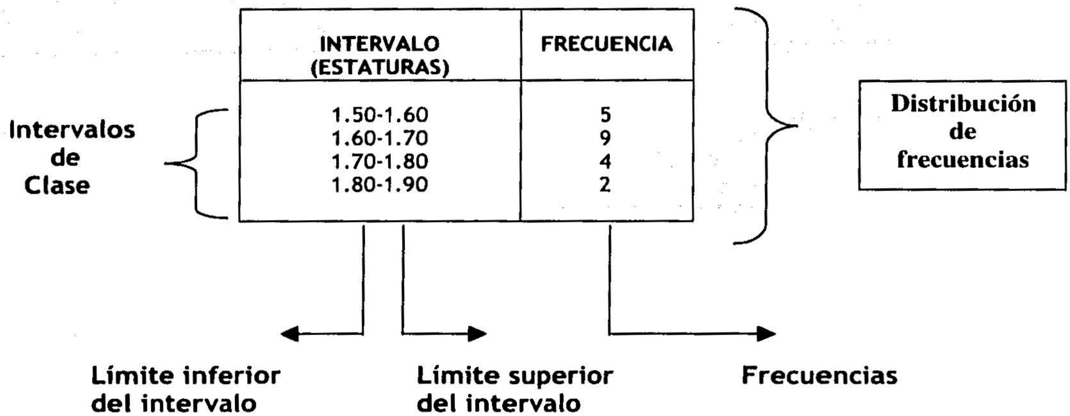
INTERVALO	FRE CUENCIA	FRECUENCIA RELATIVA	FRECUENCIA ACUMULADA	FRECUENCIA REL. ACUM.	MARCA DE CLASE
46-54	1	.066	1	.066	50
54-62	0	.000	1	.066	58
62-70	1	.066	2	.132	66
70-78	2	.133	4	.265	74
78-86	4	.266	8	.531	82
86-94	7	.466	15	.997	90

Si observamos los intervalos 62-70 y 70-78 se cumple la condición (a,b], en donde los datos que son iguales a 70 solo son tomados en cuenta en el tercer intervalo, lo cual impide que se dupliquen estos datos.

El presentar nuestra variable estadística de ésta manera nos permite hacer las siguientes observaciones:

- Solo un aspirante obtuvo un puntaje reprobatorio, siendo esto el 6.6 % de la muestra.
- El 46.6% de la muestra obtuvo un buen puntaje de 86 a 94 aciertos.

Una vez vista la manera de cómo se realiza una tabla de distribución de frecuencias por intervalos mostraremos las partes que la integran en el siguiente ejemplo:



2.6.4 REGLAS PARA FORMAR UNA DISTRIBUCIÓN DE FRECUENCIAS.

En el ejemplo anterior se mostró como construir una distribución de frecuencias para los datos sobre un examen y explicamos como se puede interpretar dicha distribución, a continuación presentaremos principios que se deben emplear para construir una distribución de frecuencias para un conjunto de datos:

1. **Determinar el número de clases.** En general es mejor tener entre 5 y 20 clases, a mayor número de datos, mayor número de clases. Si el número de clases es muy pequeño, podríamos estar encubriendo, características importantes de los datos, debido al agrupamiento. Si tenemos demasiadas clases, pueden aparecer clases vacías y la distribución perdería su significado. El número de clases debe determinarse a partir de la cantidad de datos presentes y de su uniformidad.

2. **Encontrar la amplitud y determinar el ancho de las clases.** Como regla general para encontrar el ancho de las clases (Tamaño del intervalo), divida la diferencia entre la observación mayor y la menor por el número de clases deseado y añádale lo necesario al cociente para obtener una cifra conveniente para el ancho de las clases. Todas las clases deben tener el mismo ancho. Esto nos permite hacer comparaciones uniformes de las frecuencias de clase.

3. **Determinar los limites de las clases.** Para determinar los límites de las clases, empiece con la clase más baja, de modo que incluya a la observación más pequeña. Entonces añada, las clases restantes.

Una vez que se han presentado estas reglas se resolverá el siguiente ejemplo para una mejor comprensión de la construcción de una distribución de frecuencias por intervalos.

Ejemplo 1: Se realizó un estudio acerca del número de empleos que ofrecen ciertas compañías a los Técnicos electricistas, en el área de mantenimiento teniéndose los siguientes resultados (Se encuestaron a 45 compañías).
Construya la Distribución de frecuencias para los datos con 5 intervalos:

6	4	3	7	9	8	11	4	12
2	5	2	3	3	7	7	5	5
4	3	2	8	9	11	10	6	4
3	5	6	8	9	15	12	11	6
7	7	8	9	2	2	8	9	10

Solución:

a) El número de intervalos ya ha sido proporcionado y será 5, con este número garantizamos que todos los datos queden dentro de un intervalo.

b) Calculando:

Rango = V_{mayor}-V_{menor} =15-2=13

Tamaño de intervalo = t =Rango÷Número de intervalos deseados

t= 13/5=2.6 Por lo que es conveniente tomar como t = 3

c) Para determinar los límites del primer intervalo, comenzaremos con el dato menor que es 2, recuerda que el intervalo debe de ser (a, b].

INTERVALOS	FRECUENCIA	FRECUENCIA RELATIVA	FRECUENCIA ACUMULADA	FRECUENCIA REL. ACUM.
2-5	18	.400	18	.400
5-8	14	.311	32	.711
8-11	10	.222	42	.933
11-14	2	.044	44	.977
14-17	1	.022	45	.999

De la anterior tabla de distribución de frecuencias podemos afirmar:

- El 40% de las compañías ofrecen entre 2 y 5 empleos para Técnicos electricistas.
- Solo tres compañías cuentan con una cartera de más de 12 plazas para el área Técnica.

2.6.5 REPRESENTACIONES GRÁFICAS : HISTOGRAMA Y POLÍGONO DE FRECUENCIAS.

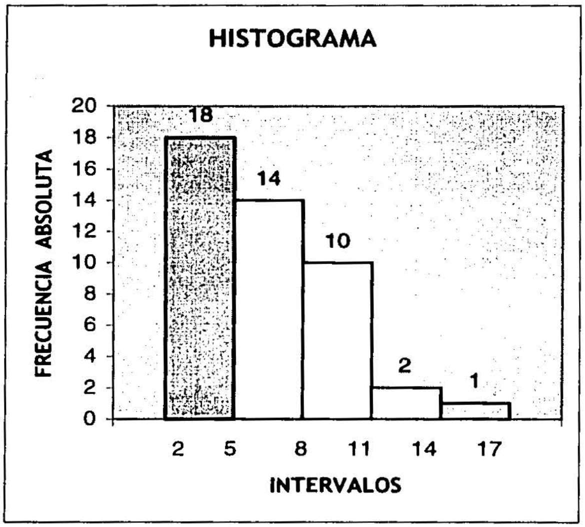
Una vez obtenida la Distribución de frecuencias de nuestros datos es muy útil contar con métodos gráficos para obtener una descripción general y rápida de los datos recolectados y para su presentación, ya que al visualizarlos nos ayudará a tener una mejor idea de lo que sucede con nuestra muestra.
El primer método que presentaremos será el Histograma.

Histograma : Es una representación gráfica de una función de distribución de frecuencias mediante rectángulos de área proporcional a la frecuencia correspondiente.
Se puede construir un histograma de frecuencias contra intervalos, frecuencias relativas acumuladas contra intervalos, etc.

Ejemplo 1: De la tabla de distribución de frecuencias del ejemplo anterior construya el histograma de frecuencias contra intervalos.

Solución : En el eje horizontal (eje de las X) se colocan los intervalos y en el eje vertical (eje de las Y) las frecuencias. Se escoge una escala que nos permita una observación clara del histograma.

INTERVALOS	FRECUENCIA
2-5	18
5-8	14
8-11	10
11-14	2
14-17	1

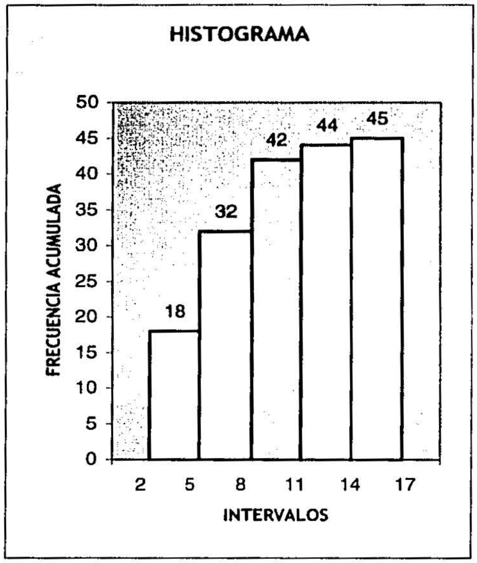


Ejemplo 2: Obtenga el histograma del ejercicio anterior de intervalos contra frecuencia acumulada.

Solución :

Las columnas a graficar son las siguientes y el histograma que se obtiene es el siguiente:

INTERVALOS	FRECUENCIA ACUMULADA
2-5	18
5-8	32
8-11	42
11-14	44
14-17	45



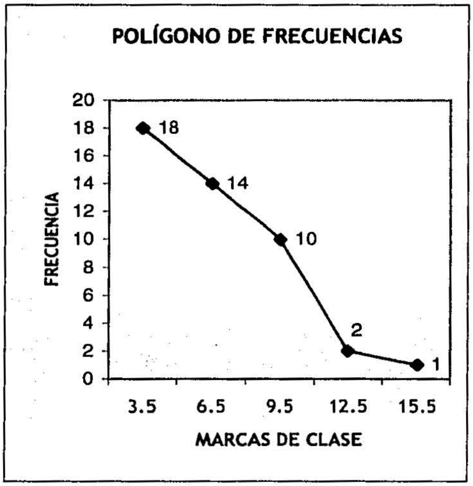
Otro método gráfico es el polígono de frecuencias, a continuación se da la definición de este así como un ejemplo.

Polígono de frecuencias: Es un gráfico de líneas a rectas trazadas, sobre los puntos medios o marcas de clase de cada rectángulo del histograma.

Ejemplo 3: Dibuje el polígono de frecuencias de Intervalos contra frecuencia del ejercicio anterior.

Solución : Para la construcción del polígono de frecuencias es necesario tener las marcas de clase de cada intervalo, que se obtiene del promedio de los límites del intervalo.

INTERVALOS	FRECUENCIA	MARCA DE CLASE
2-5	18	3.5
5-8	14	6.5
8-11	10	9.5
11-14	2	12.5
14-17	1	15.5

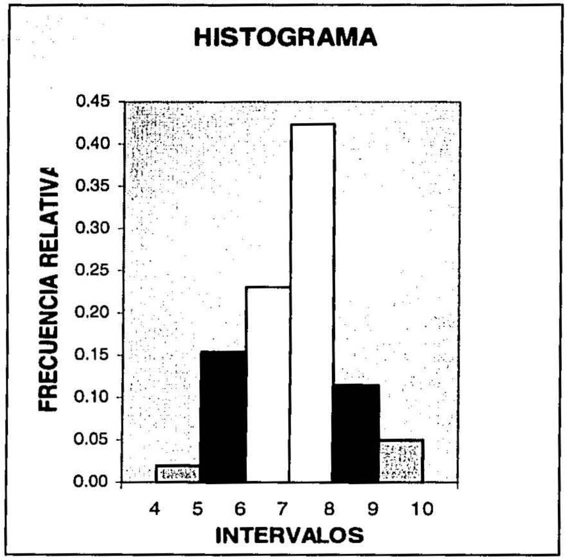


Ejemplo 4: De la siguiente Distribución de frecuencias obtenga el histograma de frecuencias de: intervalos contra frecuencia relativa y el polígono de: intervalos contra frecuencia acumulada.

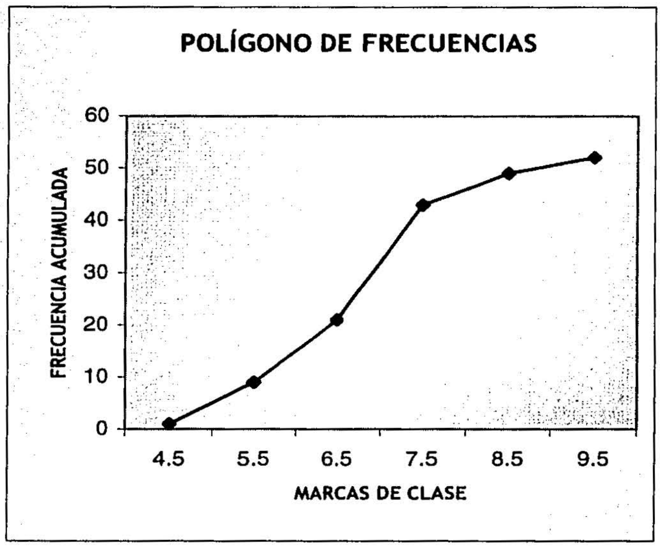
La siguiente tabla muestra los datos de los promedios obtenidos de los alumnos de 6° semestre de Bachillerato:

INTERVALO	FRECUENCIA	FRECUENCIA RELATIVA	FRECUENCIA ACUMULADA	FRECUENCIA ACUM. REL.	MARCA DE CLASE
4-5	1	0.019	1	0.019	4.5
5-6	8	0.154	9	0.173	5.5
6-7	12	0.231	21	0.403	6.5
7-8	22	0.423	43	0.826	7.5
8-9	6	0.115	49	0.942	8.5
9-10	3	0.058	52	1.000	9.5

El histograma de frecuencias de: intervalos contra frecuencia relativa queda de la siguiente manera:



El polígono de intervalos contra frecuencia acumulada se muestra a continuación:



Como se observa la gráfica sale ascendente ya que se graficó la frecuencia acumulada.

CAPÍTULO III

MEDIDAS DE TENDENCIA CENTRAL

Para poder obtener consecuencias y deducciones validas de los datos de una estadística, es muy útil contar con información sobre los valores al centro y sobre las distancias que están unos valores con respecto de otros. Las primeras medidas se llaman medidas de tendencia central, las cuales se tratarán en este capitulo. Comenzaremos con Notación sumatoria que nos ayudará a comprender mejor el tema, así como el desarrollo de las fórmulas.

3.1 NOTACIÓN FUNCIONAL.

La suma es una de las operaciones que se utiliza con mayor frecuencia en el análisis de datos estadísticos, por lo cual es necesario un símbolo para indicar ésta operación entre los datos de la muestra. Este símbolo es llamado sumatoria. Para comprender la notación de sumatoria se requiere estar familiarizado con dos temas del Álgebra elemental: **La notación funcional y la noción de sucesión**. La notación funcional es importante porque a menudo necesitaremos sumar una función de las mediciones muestrales, por ejemplo sus cuadrados. Así la función de las mediciones muestrales que se va a sumar será dada como una fórmula en **notación funcional**.
Comenzaremos por dar la definición de función, para comprender mejor lo que es **notación funcional**.

Función: Una función consiste en dos conjuntos de elementos y una correspondencia definida entre un elemento del primer conjunto digamos x, y un elemento del segundo conjunto, digamos y, tal que a cada elemento de x le corresponda uno y solo un elemento de y.
Una vez dada la definición podríamos entender que una función estará dada por una pareja identificada como (x, y)

Por ejemplo: Si quisiéramos calcular el valor de $Y = x+3$ para $x=1$, $x=2$ y $x=3$

Solución:

X	Y
1	4
2	5
3	6

La notación funcional es útil y ventajosa cuando deseamos indicar el valor de la función y para un valor especifico de x .
A continuación se presenta dos ejemplos, uno de ellos tiene la particularidad de asignar una valor **C** llamado constante, al calcular la función.

Ejemplo 1: Dada la función $Y = \frac{1}{x}$, donde $x \neq 0$.

Encuentra $Y(3)$

$$Y(3) = \frac{1}{3}$$

Encuentra $Y(6)$

$$Y(6) = \frac{1}{6}$$

Ejemplo 2: Sea la función $g(x) = 7$

Encuentra $g(4)$

$$g(4) = 7$$

Encuentra $g(5)$

$$g(5) = 7$$

Observa que $g(x)$ siempre es igual a 7, para cualquier valor de x puesto que el valor de ésta función para cualquier valor asignado a x será 7 (la constante 7).

3.2 SUCESIONES NUMÉRICAS.

Al realizar un estudio acerca de una variable aleatoria, tendremos un conjunto de mediciones, de las cuales una ocupará el primer lugar, otra el segundo y así sucesivamente, empíricamente entendemos lo que es una sucesión pero daremos una definición más formal.

Sucesión : Una sucesión es un conjunto de objetos a_1 , a_2 , a_3 ,a_4. ,ordenados en el sentido de que podamos identificar al primer elemento del conjunto, es decir a_1, al segundo a_2, etc.

Casi siempre los elementos de las sucesiones son números. Puesto que los elementos de una sucesión están ordenados, es natural tratar de encontrar una fórmula para un elemento cualquiera de la sucesión **en función de su posición.**
Ejemplo 1: De la siguiente sucesión de números 3,4,5,6,7,....., encuentre la fórmula que la determina.

Solución: Si observamos la sucesión detenidamente se encuentra que cada elemento es una unidad mayor que el elemento precedente, por lo que podemos asignar una variable **y** para cada posición de la sucesión, de tal forma que **y** toma los valores de 1,2,3,4 para las posiciones 1,2,3,4 respectivamente.

Por lo que la fórmula está dada por $f(y) = y + 2$

Calculando la función para los valores anteriores

y	f(y)
1	3
2	4
3	5
4	6
5	7

Ejemplo 2: Dada la siguiente sucesión 1,4,9,16,25 ...,encuentre la fórmula.

Solución: Al analizar la sucesión se observa que cada elemento fue el resultado de elevarlo al cuadrado.

y	f(y)
1	1
2	4
3	9
4	16
5	25

Entonces la función es $f(y) = y^2$

Ejemplo 3: Dada la siguiente sucesión 1,8,27,64,... ,encuentre la fórmula.

Solución: observando la sucesión se ve que cada número fue elevado al cubo para obtener la sucesión. Por lo que la formula es:

$$f(y) = y^3$$

3.3 NOTACIÓN SUMATORIA.

Una vez visto los dos temas anteriores podemos hacer uso de ellos para comenzar con la explicación de **Notación sumatoria.** Al obtener nuestros datos en el muestreo recordamos que cada uno de ellos ocupa un lugar en la sucesión, por lo cual podemos calcular, por ejemplo el cuadrado de los primeros 5 términos de la sucesión, mediante el uso de $\sum$ (sigma mayúscula), que corresponde a la letra S de suma, la cual nos indica que debemos sumar en este caso los primeros cinco elemento de la sumatoria previamente elevados al cuadrado. Mediante el uso de la sumatoria quedaría de la siguiente forma:

$$\sum_{y=1}^{5} y^2$$

Realizaremos algunos ejemplos para comprender mejor el concepto de **Sumatoria.**

Ejemplo1: Calcula la siguiente sumatoria.

$$\sum_{y=1}^{5} y^2 = (1)^2 + (2)^2 + (3)^2 + (4)^2 + (5)^2$$
$$= 1 + 4 + 9 + 16 + 25$$
$$= 55$$

Ejemplo2: Obtenga el resultado de la siguiente sumatoria $\sum_{x=2}^{4} 5x$

$$\sum_{x=2}^{4} 5x = 5(2)+5(3)+5(4)$$
$$= 10+15+20$$
$$= 45$$

Nota: El elemento típico es función solo de la variable de adición. Todos los otros símbolos se consideran como constantes

A continuación daremos algunos teoremas acerca de las sumatorias, para una mejor comprensión del tema.

Teorema 1.- Sea C una constante (es decir un elemento que no depende de la variable de adición) y sea y la variable de adición. Entonces:

$$\sum_{y=1}^{n} c = nc$$

Demostración:

$$\sum_{y=1}^{n} c = c + c + c + c + \ldots + c$$

donde la suma contiene n elementos. Entonces:

$$\sum_{y=1}^{n} c = nc$$

Ejemplo 3: Obtenga la siguiente sumatoria $\sum_{y=1}^{3} 5a$. Observa que y es la variable de adición y, por tanto, a es una constante.

$$\sum_{y=1}^{3} 5a = 3(5a)$$
$$= 15a$$

Teorema 2.- Sea C una constante. Entonces

$$\sum_{i=1}^{n} cy_i = c\sum_{i=1}^{n} y_i$$

Demostración:

$$\sum_{i=1}^{n} cy_i = cy_1 + cy_2 + cy_3 + \ldots + cy_n$$
$$= c(y_1 + y_2 + y_3 + \ldots + y_n)$$
$$= c\sum_{i=1}^{n} y_i$$

Ejemplo 4: Calcula la siguiente sumatoria $\sum_{x=2}^{4} 5x$

$$\sum_{x=2}^{4} 5x = 5(2)+5(3)+5(4)$$
$$= 5(2+3+4)$$
$$= 5\sum_{x=2}^{4} x$$

Teorema 3.- $\sum_{i=1}^{n}(x_i + y_i + z_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} z_i$

Demostración:

$$\sum_{i=1}^{n}(x_i + y_i + z_i) = x_1 + y_1 + z_1 + x_2 + y_2 + z_2 + x_3 + y_3 + z_3 + \ldots.. + x_n + y_n + z_n$$

Reagrupando términos, tenemos:

$$= \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} z_i$$

Ejemplo5: Calcule la siguiente sumatoria $\sum_{x=1}^{3}(x^3 + ax + 7)$

Solución:

$$\sum_{x=1}^{3}(x^3 + ax + 7) = \sum_{x=1}^{3} x^3 + \sum_{x=1}^{3} ax + \sum_{x=1}^{3} 7$$
$$= \sum_{x=1}^{3} x^3 + a\sum_{x=1}^{3} x + 3(7)$$
$$= (1 + 8 + 27) + a(1 + 2 + 3) + 21$$
$$= 6a + 57$$

Ejemplo 6: Calcula la siguiente sumatoria $\sum_{x=2}^{5}(x-1)$

Solución:

$$\sum_{x=2}^{5}(x-1) = (2-1) + (3-1) + (4-1) + (5-1)$$
$$=10$$

Ejemplo 7: Obtenga $\sum_{x=2}^{5} x - 1$

Solución:

$$\sum_{x=2}^{5} x - 1 = (2 + 3 + 4 + 5) - 1$$
$$=13$$

Observa la diferencia entre los dos ejercicios anteriores, el solo hecho de colocar paréntesis los hace diferentes.

3.4 MEDIDAS DE TENDENCIA CENTRAL.

Los métodos gráficos nos ayudan a tener una visión de lo que sucede con nuestros datos, pero si quisiéramos más información acerca de estos, recurrimos al uso de métodos numéricos para poder obtener consecuencias y deducciones válidas, lo cual se logra por medio del uso de las medidas de tendencia central que nos proporcionan información sobre los valores al centro, así como la tendencia de agrupación de los datos. Estas medidas de tendencia central son: **Media, Mediana y Moda.**

3.4.1 MEDIA.

La Media es una medida de tendencia central equivalente al promedio aritmético de todos los valores de la variable que están en la muestra.
En Estadística una media se representa por medio de $\overline{X}$ (Equis barra).

Para su cálculo es necesario considerar primero la manera en que se tienen presentados los datos (ordenados y agrupados por intervalos etc.)

Media para datos no agrupados ni ordenados

Considere el siguiente grupo de datos de 15 alumnos que corresponde al promedio obtenido al finalizar la secundaria.

7.1	7.4	8.2	6.8	8.9
7.4	6.8	7.7	7.9	6.7
8.5	9.1	7.4	7.6	8.6

La fórmula para calcular la media de ésta muestra es:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ , es decir } \bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}$$

Donde :

n: Es el número de datos.
x: Es la variable.

Calculando la Media para estos datos se tendrá:

$$\bar{x} = \frac{7.1+7.4+8.2+6.8++8.9++7.4+6.8++7.7+7.9+6.7+8.5+9.1+7.4+7.6+8.6}{15}$$

$$\bar{x} = \frac{116.1}{15} = 7.74$$

Este resultado nos dice que el promedio obtenido por los alumnos al egresar de la secundaria es de 7.74

Media para datos agrupados por frecuencia

La siguiente tabla muestra las edades de 20 alumnos que ingresaron a nivel bachillerato, obtenga la media.

VARIABLE x_i	**FRECUENCIA** f_i
14	4
15	5
16	3
17	4
19	2
22	1
25	1

La Media para datos agrupados por frecuencia se calcula de la siguiente manera:

$$\bar{x} = \frac{1}{n}(f_1x_1 + f_2x_2 + f_3x_3 + \ldots + f_nx_n) = \frac{1}{n}\sum_{i=1}^{n} f_ix_i$$

Donde :

n: Es el número de datos.
f_i: Es la frecuencia absoluta.
x: Es la variable

Calculando la media se obtiene:

$$\bar{x} = \frac{(14)(4)+(15)(5)+(16)(3)+(17)(4)+(19)(2)+(22)(1)+(25)(1)}{20}$$

$$\bar{x} = \frac{56+75+48+68+38+22+25}{20}$$

$$\bar{x} = \frac{332}{20}$$

$$\bar{x} = 16.6$$

Por lo que la edad promedio al entrar al bachillerato de esta muestra es de 16.6 años.

Media para datos ordenados y agrupados por intervalos.

La fórmula para obtener la media para datos ordenados y agrupados por intervalos es la siguiente:

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

Donde:

n: Es el número total de datos.
f_i : Frecuencia absoluta.
C_i : Marca de clase.
h: Es el número de intervalos.

Los siguientes datos son los salarios percibidos mensualmente y son presentados en la siguiente distribución de frecuencias por intervalos de la cual obtendremos su media.

INTERVALOS	FRECUENCIA f_i	MARCA DE CLASE C_i
$ 2000-3500	6	2750
3500-5000	5	4250
5000-6500	4	5750
6500-8000	2	7250
8000-9500	3	8750
9500-11000	2	10250

Recuerda que la marca de clase se obtiene del promedio aritmético de los intervalos de cada clase.

Calculando la Media:

$$\bar{x} = \frac{(2750)(6)+(4250)(5)+(5750)(4)+(7250)(2)+(8750)(3)+(10250)(2)}{26}$$

$$\bar{x} = \frac{16500+21250+23000+14500+26250+20500}{26}$$

$$\bar{x} = \frac{122000}{26}$$

$$\bar{x} = 4692.3$$

La media nos dice que el salario mensual promedio es $ 4,692.30

3.4.2 MEDIANA.

Otra de las medidas de tendencia central es la **Mediana**, que es el valor del término central de toda una serie de **datos ordenados** conforme a los valores ya sea en forma creciente o decreciente. La mitad de las observaciones debe tener valor menor al de la mediana y la otra mitad debe tener valor mayor que el de la mediana.

Por lo tanto la mediana solo se puede obtener de datos ordenados. Para simbolizar a la mediana comúnmente se utiliza Md. Calculemos la mediana de la siguiente serie de datos previamente ordenados:

2 , 6 , 9 , 13 , 17 , 21 , 27

La mediana de ésta serie es el valor del cuarto término, o sea 13, puesto que antes y después se encuentra un mismo número de términos.

Cálculo del número de orden de la mediana.

Como la mediana es el valor del término que ocupa el punto medio de una serie ordenada, se deduce que si al número de orden de la mediana se le suma el número de términos que le siguen o que le anteceden, se obtiene el número total de términos de la serie (N). Ahora bien, como el número de términos que anteceden ala mediana es igual a su número de orden menos uno, resulta, que sí representamos por N_0 al número de orden de la mediana, se tiene que:

$$N_0 + (N_0 - 1) = N \quad \text{de donde}$$
$$N_0 + N_0 - 1 = N$$
$$2N_0 = N + 1 \quad \text{luego entonces}$$
$$N_0 = \frac{N+1}{2}$$

Esta igualdad nos dice que el número de orden o localización de la mediana de una serie ordenada es igual al número de términos de la serie más uno dividido entre dos.

Mediana para una serie impar.

Si en la fórmula, suponemos que N sea un número impar el cociente $\frac{N+1}{2}$ resulta entero. Por ejemplo calculando la mediana para la siguiente serie de datos:

1 , 4 , 5 , 8 , 10 , 12 , 14 , 18 , 27 , 33 , 35

La serie ya ésta ordenada entonces calculando el valor que debe tener la mediana en la serie:

$N_0 = \frac{11+1}{2} = 6$ ∴ el lugar que ocupa la mediana es el sexto que corresponde al valor de 12 entonces:

Md= 12

Mediana para una serie par.

Suponiendo en la igualada que N tenga un valor par, el cociente $\frac{N+1}{2}$ resulta fraccionaria, por lo que en tal caso se acostumbra tomar como valor de la mediana, el promedio aritmético del término anterior y posterior del número de orden calculado para la mediana.

De la siguiente serie calcule la mediana:

$$3\ ,\ 5\ ,\ 6\ ,\ 7\ ,\ 8\ ,\underbrace{\ 9\ ,\ 12\ }\ ,\ 13\ ,\ 14\ ,\ 18\ ,\ 22\ ,\ 24$$

Recuerda que la mediana divide en dos conjuntos de igual cardinalidad y siendo el número de términos doce y calculando el número de orden de la mediana:

$$N_0 = \frac{12+1}{2} = 6.5$$

El resultado anterior nos dice que el número de orden de la mediana en la serie es 6.5. El término de rango 6.5 no existe. Entonces se calcula el promedio aritmético del término anterior y posterior del número de orden calculado para la mediana.

$$Md = \frac{9+12}{2} = \frac{21}{2} = 10.5$$

Mediana para una distribución de frecuencias.

Como en el fondo, toda serie de frecuencias es una serie simple en la que ésta simplificada la escritura de los términos que en la serie simple se repiten, el cálculo de la mediana de una serie de frecuencias es igual al de la serie simple. De tal forma que podemos decir que: El número de orden que ocupa la mediana en una serie de frecuencias es igual a la suma de frecuencias más uno, dividida entre dos.

Ejemplo 1: Calcule la mediana para la siguiente distribución de frecuencias, los datos son acerca del salario por hora obtenido por 55 técnicos electricistas.

SALARIO POR HORA	FRECUENCIA
$ 100	4
150	9
200	25
250	13
300	3
350	1

Solución : Primero calculamos el número de orden de la mediana

$$N_0 = \frac{55+1}{2} = 28$$

El resultado anterior indica que el valor del término de número de orden 28 es el que corresponde a la mediana.
Ahora hay que localizar este valor en la distribución de frecuencias para ver a que salario corresponde. Los 4 Técnicos que ganan $100 y los 9 que tienen un salario de $150 suman 13 términos, faltando por lo tanto 15 términos para llegar al lugar que ocupa la mediana. este lugar se encuentra entre los 25 Técnicos que tienen un salario de $200 es decir:

$$Md = \$200$$

Ejemplo 2: De la siguiente distribución de frecuencias calcule la mediana.

SALARIO POR HORA	FRECUENCIA
$ 100	4
150	10
200	13
250	15
300	8
350	4

Si observas el número de datos registrados en esta distribución es par, por lo que calculándole número de orden que ocupa la mediana tendremos:

$$N_0 = \frac{54+1}{2} = 27.5$$

Este resultado indica que la mediana de la serie tiene un valor cualquiera comprendido entre los valores de los términos que ocupan los lugares 27 y 28. El valor del término de número de orden 27 es de $200 (Ya que 4+10+13=27) y el 28 es de $250 (Ya que 4+10+13+15=42 y el dato 28 se encuentra ahí), tomando como valor de la mediana el promedio aritmético de estos valores resulta:

$$Md = \frac{200+250}{2} = 225$$
$$Md = \$225$$

Mediana para una distribución de frecuencias por intervalos.

Cuando los datos se presentan por medio de una distribución de frecuencias por intervalos, la mediana sigue siendo el valor del elemento medio, sin embargo para el cálculo de la Md se utiliza la siguiente fórmula:

$$Md = L_i + \left[\frac{n/2 - \left(\sum f_i\right)}{f..mediana}\right]C$$

En donde:

L_i = *Límite inferior del intervalo de clase donde se encuentra la mediana.*

n = *Número de datos.*

$\sum f_i$ = *Suma de todas las frecuencias anteriores a la clase en donde se encuentra la mediana.*

$f.mediana$ = *Frecuencia de la clase de la mediana.*

C = *Tamaño del intervalo dela clase de la mediana.*

Ejemplo 3: Calcule la mediana de la siguiente distribución de frecuencias por intervalos.

INTERVALOS	FRECUENCIA
$1200-1300	3
1300-1400	5
1400-1500	9
1500-1600	12
1600-1700	5
1700-1800	4
1800-1900	2

Puesto que la mediana es el valor central y la muestra es par, para localizar en que intervalo se trabajará se hacen los cálculos y el dato 20 y 21 se encuentran en el cuarto intervalo que será por tanto la clase de la mediana.

$L_i = 1500$
$n = 40$
$\sum f_i = 3+5+9=17$
$f.mediana = 12$
$C = 100$

Sustituyendo los valores en la fórmula:

$$Md = 1500 + \left[\frac{40/2 - 17}{12}\right]100$$
$$Md = 1500 + \left[\frac{20-17}{12}\right]100$$
$$Md = 1500 + \left[\frac{3}{12}\right]100$$
$$Md = 1500 + \left[\frac{300}{12}\right]$$
$$Md = 1500 + 25 = 1525$$
$$Md = \$1525$$

3.4.3 MODA

La moda es otra de las medidas de tendencia central también llamada modo, valor normal o valor de máxima frecuencia. La moda es el valor que se presenta con mayor frecuencia en la muestra.
Ciertas distribuciones presentan más de una considerando moda como el o los máximos valores del polígono de frecuencias (plurimodales). Para identificar el valor de la moda lo haremos por medio de Mo.

Moda para datos ordenados.

De los siguientes datos ordenados obtenga la moda:

2 , 2 , 2 , 4 , 4 , 5 , 5 , 7 , 7 , 7 , 7 , 11 , 11 , 15

Siendo que la moda es el dato que más se repite en la muestra es muy sencillo obtenerla de la anterior. Siendo Mo = 7

Moda para una distribución de frecuencias.

Ya teniendo los datos ordenados por medio de una distribución de frecuencias es más fácil encontrar la moda. Calculando la moda para la siguiente distribución de frecuencias:

SALARIO POR HORA	FRECUENCIA
\$ 100	4
150	9
200	25
250	13
300	3
350	1

Si observamos la distribución el dato que presenta mayor número de frecuencia es el que corresponde al de \$200.

Entonces Mo = \$200

Moda para una distribución de frecuencias por intervalos.

Para el cálculo de la moda en una distribución de frecuencias por intervalos recurriremos al uso de la siguiente fórmula:

$$Mo = L_i + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right] \cdot C$$

En donde:

$Mo = Moda$.
L_i = *Límite inferior del intervalo donde se encuentra la moda.*
Δ_1 = *Diferencia de la frecuencia del intervalo donde se encuentra la moda y la frecuencia anterior.*
Δ_2 = *Diferencia de la frecuencia del intervalo donde se encuentra la moda y la frecuencia posterior.*
C = Tamaño del intervalo.

Aplicaremos la fórmula en el siguiente ejercicio. Obtenga la moda de la siguiente distribución de frecuencias por intervalos.

INTERVALOS	FRECUENCIA	MARCA DE CLASE
4-5	2	4.5
5-6	8	5.5
6-7	12	6.5
7-8	25	7.5
8-9	6	8.5
9-10	7	9.5

Al observar la distribución de frecuencias, el intervalo de 7-8 es el que cuenta con mayor número de frecuencias por lo tanto se trabajará en este intervalo.

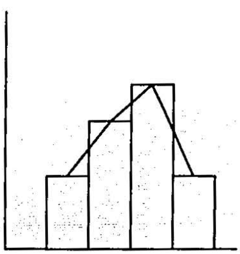
$$L_i = 7$$
$$\Delta_1 = 25 - 12 = 13$$
$$\Delta_2 = 25 - 6 = 19$$
$$C = 1$$

Sustituyendo los valores en la fórmula:

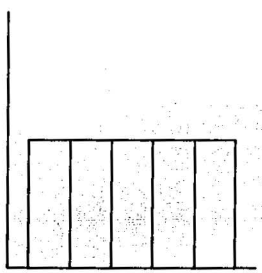
$$Mo = 7 + \left(\frac{13}{13+19} \right) \cdot 1 = 7 + \left(\frac{13}{32} \right) \cdot 1$$
$$Mo = 7 + .40625 = 7.40625$$

Muestras y sus modas.

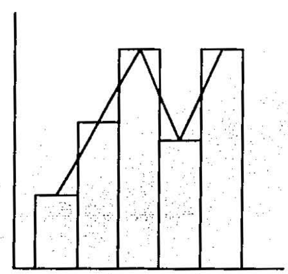
Las siguientes ilustraciones de las muestras y su o sus modas se ilustran a continuación:



Una Moda



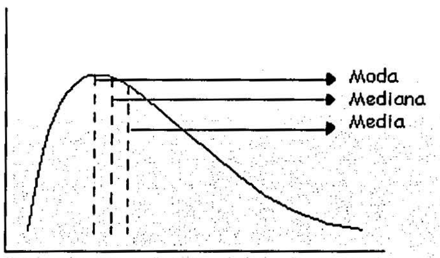
Sin Moda



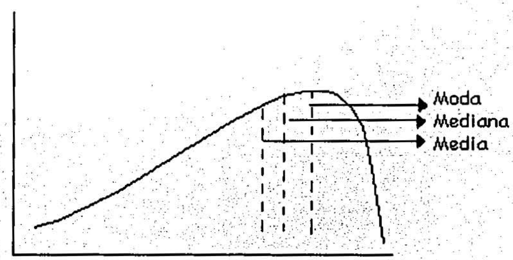
Bimodal

3.4.4 RELACIÓN EMPÍRICA ENTRE MEDIA, MEDIANA Y MODA.

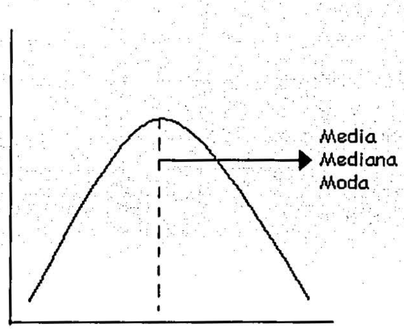
En toda distribución simétrica la Media, Mediana y Moda coinciden en valor. En una distribución asimétrica positiva, la media siempre es mayor que la mediana. En una distribución asimétrica negativa , la media siempre es menor que la mediana. Estas dos últimas relaciones siempre son verdaderas. Lo anterior se ilustra a continuación gráficamente.



Asimétrica positiva



Asimétrica negativa



Simétrica

Considerando el uso de estas medidas de promedio para la representación de datos de la población, el valor de la moda indica la posición de la mayoría de los valores observados, puede ser útil como medida descriptiva de un grupo de la población, aunque solo si existe una moda claramente perceptible, por otra parte la mediana, es siempre una medida excelente para representar el nivel típico de los valores observados, como los índices salariales de una población.

Esto es muy independiente de la existencia de una o más modas o de que la población sea simétrica o asimétrica. También la media aritmética es excelente como valor representativo de una población, aunque solo si la población es claramente simétrica. Así la mediana es por lo general una mejor medida de posición de datos para la descripción de datos de la población.

Por otro lado considerando el uso de las tres medidas de posición en relación con datos muestrales, se recordará que el propósito de la inferencia estadística con datos muestrales es producir enunciados de probabilidad sobre la población de la que fue sustraída la muestra. La moda no es una medida aceptable de posición respecto de datos muestrales, porque su valor puede variar ampliamente de una muestra a otra, por otro lado la mediana es mejor que la moda porque su valor es más estable. **No obstante el valor de la media es el valor más estable de estas tres medidas.**

CAPÍTULO IV

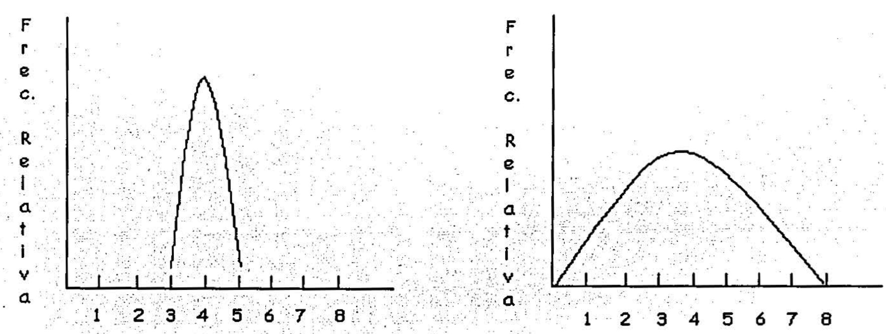
MEDIDAS DE DISPERSIÓN.

4.1 INTRODUCCIÓN.

Una vez localizado el centro de una distribución de datos, nuestro siguiente paso es proporcionar una medida de la variabilidad o dispersión de estos. Las medidas de dispersión son las que nos ayudan a encontrar lo distanciados que están unos valores con respecto de otros.

Esta medidas son: **la Varianza, la Desviación Estándar y la Desviación Media.**

Para una mejor comprensión de esto, consideremos las siguientes distribuciones de la siguiente figura.



Variabilidad o dispersión de los datos

Las dos distribuciones tienen un centro ubicado en x=4, pero hay una gran diferencia en la variabilidad de las observaciones alrededor de la media para las dos distribuciones, la mayoría de las observaciones en la primer figura varían de 3 a 5, mientras que en la segunda lo hacen de 0 a 8. La variación es una característica muy importante de los datos. Por ejemplo si fabricamos tornillos, una variación excesiva en el diámetro de los tornillos implicaría un alto porcentaje de producción defectuosa.

Por otro lado si se utiliza en un examen para hacer una selección de los mejores empleados, sería muy poco satisfactorio que el examen produjera una variabilidad pequeña, ya que esto complicaría la selección.

Además de la importancia práctica de la variación en los datos, es evidente, que una medida de esta característica es necesaria para construir una imagen mental de la distribución de frecuencias.

4.2 DESVIACIÓN MEDIA.

Hasta fines del siglo pasado, la desviación media fue la medida de dispersión de más uso, su desplazamiento del arsenal estadístico se debió a la aparición del concepto de Desviación estándar, otra medida de dispersión con mejores propiedades algebraicas y que genera valores numéricos muy parecidos a los que se obtienen con la desviación media. Aunque haya caído en desuso, conviene estudiarla debido a que su significado " fácil de comprender ", facilita la comprensión de la desviación estándar, concepto de gran utilidad y con el cual está muy relacionada.

Comenzaremos por explicar lo que es **desviación:** llamaremos desviación de una variable a la diferencia que tiene su valor con respecto a la media aritmética de la muestra.
Desviación o desvío: $x_i - \bar{x}$

Suma de los desvíos o desviaciones.

La suma de los desvíos será: $D = \sum_{i=1}^{n} (x_i - \bar{x})$

Con el siguiente ejemplo se aclarará el concepto de desvío: Considere la siguiente muestra:

1,2,5,6,7,10,11,12

Lo primero que se tiene que hacer es calcular la Media.

$$\bar{x} = \frac{1+2+5+6+7+10+11+12}{8} = \frac{54}{8} = 6.75$$

Ahora elaboraremos una tabla para calcular los desvíos.

VARIABLE	DESVÍO	VALOR ABSOLUTO DEL DESVIO
1	-5.75	5.75
2	-4.75	4.75
5	-1.75	1.75
6	-.75	.75
7	.25	.25
10	3.25	3.25
11	4.25	4.25
12	5.25	5.25
TOTALES	0	26

Recuerda que el valor absoluto es: El valor numérico de una cantidad sin tomar en cuenta su signo.

La suma de los desvíos vale 0

Observando los valores obtenidos en la tabla, sugiere que la suma de los desvíos nos dará como resultado 0.

La demostración general queda como sigue:

Se trata de hallar el valor de $\sum_{i=1}^{n}(x_i - \bar{x})$

O sea de: $\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x}$ de la fórmula $\bar{x} = \frac{i}{n}\sum_{i=1}^{n} x_i$ (**)

Se deduce $\bar{x}n = \sum_{i=1}^{n} x_i$ y como por otra parte $\sum_{i=1}^{n} \bar{x} = (\bar{x} + \bar{x} + \ldots + \bar{x}...n..veces) = n\bar{x}$

Y se tiene (**) $n\bar{x} - n\bar{x} = 0$

La Desviación Media: Es la desviación promedio de los valores absolutos de las desviaciones de los datos de una variable con respecto a su media. Teniéndose entonces la siguiente fórmula para calcularla:

Desviación Media para datos sin agrupar.

La fórmula es la siguiente $$Dm = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

Para hallar la desviación media de una serie de datos sin frecuencia asociada, sin distribución de frecuencia. Basta con seguir tres pasos:

1.-Calcular su media.
2.-Se resta la media de cada dato de la variable, lo cual produce la separación de cada dato con respecto a la media, y por último
3.- Se divide la sumatoria de los valores absolutos de esas separaciones entre el total de datos.

Ejemplo: Los siguientes datos son acerca del número de unidades vendidas de ciertos circuitos integrados durante un mes. Calcule la desviación media:

X = 5,8,8,11,11,11,14,16

Calculando la media de la muestra

$$\bar{x} = \frac{5+8+8+11+11+11+14+16}{8}$$

$$\bar{x} = \frac{84}{8} = 10.5$$

VARIABLE	$x-\bar{x}$	$\lvert x-\bar{x}\rvert$
5	-5.5	5.5
8	-2.5	2.5
8	-2.5	2.5
11	0.5	0.5
11	0.5	0.5
11	0.5	0.5
14	3.5	3.5
16	5.5	5.5
	Total	21

Sustituyendo en la fórmula:

$$Dm = \frac{21}{8} = 2.625 \cong 2.6 \text{ unidades.}$$

Concluyendo podemos decir que en promedio la venta de unidades de circuitos, difiere en 2.6 unidades con respecto a la media grupal, en cualquier dirección.

Desviación Media para datos agrupados por intervalos.

La fórmula que se utiliza para el cálculo de la desviación media cuando los datos están agrupados por intervalos se transforma en la siguiente:

$$Dm = \frac{\sum_{i=1}^{n} f_i \lvert x-\bar{x}\rvert}{n}$$

Aplicaremos la fórmula a la siguiente distribución de frecuencias por intervalos:

INTERVALOS	FRECUENCIA
4-6	6
6-8	9
8-10	12
10-12	16
12-14	13
14-16	14
16-18	18
18-20	4
20-22	2

Al tener nuestros datos en una distribución de frecuencias por intervalos, se pierde la información precisa. Por ejemplo en el primer intervalo se cuenta con una frecuencia de 6, pero no sabemos cuantos de esos 6 datos son 4, 5 ó 6, por lo que es necesario el uso del promedio de los intervalos (marca de clase), para el cálculo de la media.

Calculando la media:

$$\bar{x} = \frac{5(6)+7(9)+9(12)+11(16)+13(13)+15(14)+17(18)+19(4)+21(2)}{94}$$

$$\bar{x} = \frac{30+63+108+176+169+210+306+76+42}{94}$$

$$\bar{x} = \frac{1180}{94} = 12.55$$

INTERVALOS	FRECUENCIA	MARCA DE CLASE	$\lvert x_i - \bar{x}\rvert$	$f_i\lvert x_i - \bar{x}\rvert$
4-6	6	5	7.55	45.30
6-8	9	7	5.55	49.95
8-10	12	9	3.55	42.60
10-12	16	11	1.55	24.80
12-14	13	13	0.45	5.85
14-16	14	15	2.45	34.30
16-18	18	17	4.45	80.10
18-20	4	19	6.45	25.80
20-22	2	21	8.45	16.90

Sustituyendo en la fórmula:

$$Dm = \frac{\sum_{i=1}^{n} f_i |x_i - \bar{x}|}{n} = \frac{325.6}{94} = 3.46$$

Es decir los datos de la variable x se desvían 3.46 unidades en promedio con respecto a su media.

4.3 VARIANZA.

La varianza se asemeja a la desviación media absoluta en que se basa en la diferencia entre cada valor del conjunto de datos y la media del grupo. Pero se distingue de ella en un muy importante aspecto: Cada Diferencia se eleva al cuadrado antes de sumarse.
En el caso de una población, la varianza se representa con V(x) o más habitualmente con la letra griega minúscula σ^2 (sigma cuadrada.) Algunos autores la representan por S^2 (Varianza muestral). La fórmula para obtenerla es la siguiente:

$$V(x) = S^2 = \frac{\sum (x - \bar{x})^2}{n}$$

Calcularemos la varianza para el siguiente conjunto de datos.

6, 7, 7, 7, 8, 9, 10, 11, 12, 15

La media de los datos es:

$$\bar{x} = \frac{92}{10} = 9.2$$

Calcularemos la varianza por medio de una tabla con tres columnas, una para los datos (x), otra para las desviaciones con respecto a su media $(x - \bar{x})$ y otra más para las desviaciones cuadráticas. Aplicándose posteriormente la fórmula antes descrita.

VARIABLE	$(x-\bar{x})$	$(x-\bar{x})^2$
6	-3.2	10.24
7	-2.2	4.84
7	-2.2	4.84
7	-2.2	4.84
8	-1.2	1.44
9	-0.2	0.04
10	0.8	0.64
11	1.8	3.24
12	2.8	7.84
15	5.8	33.64
	$\sum(x-\bar{x})=0$	$\sum(x-\bar{x})^2=71.6$

Ahora sustituyendo los valores obtenidos en la fórmula de la varianza:

$$S^2 = \frac{\sum(x-\bar{x})^2}{n} = \frac{71.6}{10} = 7.16$$

En general es difícil la interpretación del significado de la varianza, porque las unidades en las que se le expresa son valores elevados al cuadrado. Debido en parte a está causa es más frecuente el uso de la Desviación Estándar también llamada típica.

4.4 DESVIACIÓN ESTÁNDAR.

Esta medida de dispersión es la más adecuada por sus propiedades algebraicas, se obtiene a partir de la raíz cuadrada de la varianza. La representación de la desviación estándar esta dada por S.La fórmula es la siguiente:

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

Retomando el ejemplo anterior y calculando la desviación estándar:

$$S = \sqrt{\frac{71.6}{10}} = \sqrt{7.6} = 2.7$$

Cuya interpretación puede escribirse así: los datos de la variable X se separan 2.7 unidades en promedio respecto a la media.

A continuación se dan las fórmulas para la varianza y desviación estándar de una población así como la varianza y desviación estándar muestrales, en estás últimas la fórmula incluye un factor de corrección, a fin de que la varianza muestral sea un estimador insesgado de la varianza de la población.

Varianza de la población $\sigma^2 = \dfrac{\sum (x-\mu)^2}{n}$

Desviación estándar de la población $\sigma = \sqrt{\dfrac{\sum (x-\mu)^2}{n}}$

Varianza de la muestra $S^2 = \dfrac{\sum (x-\bar{x})^2}{n-1}$

Desviación estándar de la muestra $S = \sqrt{\dfrac{\sum (x-\bar{x})^2}{n-1}}$

Obtención de la varianza y la desviación estándar a partir de un método abreviado.

El procedimiento común para la obtención de la varianza y la desviación estándar se tornan muy laboriosas cuando la serie de datos es muy numerosa. Es necesario entonces contar con alternativas que reduzcan notoriamente el cálculo de estas medidas de dispersión. A continuación se da una explicación detallada de la obtención de estas fórmulas.

Por definición se sabe:

$$S^2 = \frac{\sum (X-\overline{X})^2}{N} \qquad (1)$$

Entonces podemos escribir

$$NS^2 = \sum (X-\overline{X})^2 \qquad (2)$$

Tomando por separado, la expresión del segundo miembro en (2) y desarrollándola.

$$\sum (X-\overline{X})^2 = \sum (X^2 - 2X\overline{X} + \overline{X}^2) \qquad (3)$$

Aplicándose reglas de la sumatoria:

$$\sum (X-\overline{X})^2 = \sum X^2 - 2\overline{X}\sum X + N\overline{X}^2 \qquad (4)$$

Por otro lado también por definición se sabe que:

$$\overline{X} = \frac{\sum X}{N} \qquad (5)$$

Entonces si sustituimos (5) en (4)

$$\sum (X-\overline{X})^2 = \sum X^2 - \frac{2\sum X}{N}\left(\sum X\right) + N\left(\frac{\sum X}{N}\right)^2$$

$$= \sum X^2 - 2\frac{\left(\sum X\right)^2}{N} + N\frac{\left(\sum X\right)^2}{N^2}$$

$$= \sum X^2 - 2\frac{\left(\sum X\right)^2}{N} + \frac{\left(\sum X\right)^2}{N}$$

$$\sum (X-\overline{X})^2 = \sum X^2 - \frac{\left(\sum X\right)^2}{N} \qquad (6)$$

Si dividimos ambos miembros de la igualdad (6) entre N, no se altera y se obtiene:

$$\frac{\sum (X-\overline{X})^2}{N} = \frac{\sum X^2 - \left(\sum X\right)^2 / N}{N}$$

Resulta que el primer miembro de la igualdad anterior, es precisamente la media de las desviaciones cuadráticas, es decir, la varianza.

Por lo tanto:

$$S^2 = \frac{\sum X^2 - \left(\sum X\right)^2 / N}{N} \qquad (7)$$

$$S = \sqrt{\frac{\sum X^2 - \left(\sum X\right)^2 / N}{N}} \qquad (8)$$

Pero estas fórmulas pueden adoptar otras formas equivalentes, igualmente útiles. si expresamos el segundo miembro de (7) como el cociente de cada uno de los términos del numerador entre el denominador común se obtendrá:

$$S^2 = \frac{\sum X^2}{N} - \frac{\frac{(\sum X)^2}{N}}{N}$$

$$S^2 = \frac{\sum X^2}{N} - \frac{(\sum X)^2}{N^2}$$

Pero el cuadrado de la sumatoria de X entre N^2, puede indicarse como el cuadrado del cociente de la sumatoria de X entre N, que no es más que la media al cuadrado.

$$S^2 = \frac{\sum X^2}{N} - \left(\frac{\sum X}{N}\right)^2$$

Por lo tanto:

$$S^2 = \frac{\sum X^2}{N} - \overline{X}^2 \qquad (9)$$

De donde:

$$S = \sqrt{\frac{\sum X^2}{N} - \overline{X}^2} \qquad (10)$$

Resumiendo **la varianza** se define como la media de las desviaciones cuadráticas y que la desviación típica, que expresa la separación promedio de los datos respecto a su media, no es mas que la raíz cuadrada de la varianza.

$$S^2 = \frac{\sum (X - \overline{X})^2}{N} \qquad S = \sqrt{\frac{\sum (X - \overline{X})^2}{N}}$$

Y por medio de pasos algebraicos obtuvimos expresiones equivalentes que reducen significativamente el trabajo de cálculo.

$$\left.\begin{aligned} S^2 &= \frac{\sum X^2 - (\sum X)^2/N}{N} = \frac{\sum X^2}{N} - \overline{X}^2 \\ S &= \sqrt{\frac{\sum X^2 - (\sum X)^2/N}{N}} = \sqrt{\frac{\sum X^2}{N} - \overline{X}^2} \end{aligned}\right\}$$

Fórmulas abreviadas de S^2 y S

Aplicaremos las fórmulas abreviadas de varianza y desviación estándar a la siguiente serie de datos.

Ejemplo 1:

X = 6 , 7 , 7 ; 7 , 8 , 9 , 10 , 11 , 12 , 15

Realizaremos la siguiente tabla, la cual solo necesita dos columnas:

X	X^2
6	36
7	49
7	49
7	49
8	64
9	81
10	100
11	121
12	144
15	225
$\sum x = 92$	$\sum x^2 = 918$

Calculando los datos para sustituirlos en las fórmulas:

$$\left(\sum X\right)^2 = 8464$$
$$N = 10$$
$$\overline{X} = 9.2$$

Sustituyendo estos valores en la fórmula (7)

$$S^2 = \frac{\sum X^2 - \left(\sum X\right)^2 / N}{N}$$

$$S^2 = \frac{918 - 8464/10}{10}$$

$$S^2 = \frac{918 - 846.4}{10}$$

$$S^2 = \frac{71.6}{10} = 7.6$$

$$S = 2.7$$

Varianza y Desviación Media para una distribución de frecuencias por intervalos.

Teniéndose una distribución de frecuencias, la fórmula para el cálculo de la varianza y la desviación estándar es la siguiente:

$$S^2 = \frac{\sum f(X - \overline{X})^2}{N} \qquad \textbf{(11)}$$

$$S = \pm\sqrt{S^2}$$

Si aplicamos el álgebra a la anterior fórmula se obtiene:

$$S^2 = \frac{\sum fX^2 - {(\sum fX)^2}/{N}}{N}$$

$$S^2 = \frac{\sum fX^2}{N} - \overline{X}^2$$

Aplicaremos estas fórmulas a dos ejemplos, uno con una distribución de frecuencias solo para datos ordenados y otra para datos ordenados y agrupados por intervalos.

Ejemplo 2: Calcular e interpretar la desviación Estándar de la distribución siguiente

Los datos son acerca del número de personas con quien comparten la habitación los alumnos del colegio y los datos arrojados son los siguientes:

VARIABLE	FRECUENCIA
0	40
1	76
2	44
3	12
4	9
5	3
6	2
7	1

La variable en ésta distribución simple de frecuencias es " número de personas ".

VARIABLE X	FRECUENCIA f	X^2	fX	fX^2
0	40	0	0	0
1	76	1	76	76
2	44	2	88	176
3	12	9	36	108
4	9	16	36	144
5	3	25	15	75
6	2	36	12	72
7	1	49	7	49
	$\sum f = 187$		$\sum fX = 270$	$\sum fX^2 = 700$

Calculando los demás valores:

$$\overline{X} = \frac{270}{187} = 1.4 \text{ personas}$$

$$\left(\sum fX\right)^2 = 72900$$

Y sustituyéndolos en la fórmula:

$$S^2 = \frac{\sum fX^2 - \left(\sum fX\right)^2 / N}{N}$$

$$S^2 = \frac{700 - 72900/187}{187}$$

$$S^2 = \frac{700 - 390}{187} = 1.66$$

$$S = \sqrt{1.66} = 1.3$$

Concluyendo : la desviación promedio de los datos de la variable X (número de personas), es 1.3 unidades con respecto a su valor medio.

Ejemplo 3: Obtenga la desviación Estándar para la siguiente distribución de frecuencias por intervalos. Los datos se refieren al tiempo dedicado al estudio, medido en horas semanarias.(En un conjunto de alumnos de 1990).

INTERVALOS	FREC.	MARCA DE CLASE (X)	X^2	fX	fX^2
1-3	50	2	4	100	200
4-6	38	5	25	190	950
7-9	26	8	64	208	1664
10-12	36	11	121	396	4356
13-15	19	14	196	266	3724
16-18	7	17	289	119	2023
19-21	7	20	400	140	2800
22-24	5	23	529	115	2645

Calculando los valores para sustituirlos en la fórmula:

$$\sum f = 188$$

$$\overline{X} = 8.2$$

$$\sum fX = 1534$$

$$\left(\sum fX\right)^2 = 2{,}353{,}156$$

$$\sum fX^2 = 18{,}362$$

Sustituyendo los valores en la fórmula:

$$S^2 = \frac{\sum fX^2 - \left(\sum fX\right)^2 / N}{N}$$

$$S^2 = \frac{18{,}362 - 2{,}353{,}156 / 188}{188}$$

$$S^2 = \frac{18{,}362 - 12{,}516.7}{188} = \frac{5{,}845.3}{188} = 31.09$$

$S = \sqrt{31.09} = 5.6$ horas semanarias.

Los datos de la variable tiempo dedicado al estudio, indagada en un conjunto de alumnos en 1990, se alejan 5.6 horas en promedio de su media aritmética.

4.5 COEFICIENTE DE VARIABILIDAD.

Tomemos las dos distribuciones propuestas a continuación y calculemos sus desviaciones estándar respectivas:

Para la distribución 1 se tiene

X	X²
10	100
15	225
20	400

En donde:

$$\sum x = 45$$
$$\overline{X} = 15$$
$$\sum X^2 = 725$$
$$\overline{X}^2 = 225$$

Calculando S:

$$S = \sqrt{\frac{\sum X^2}{N} - \overline{X}^2} \qquad S = \sqrt{\frac{725}{3} - 225} = 4.1$$

Para la distribución 2 tenemos:

X	X²
50	2500
55	3025
60	3600

Realizando los cálculos y obteniendo S:

$$\sum X = 165$$
$$\overline{X} = 55$$
$$\sum X^2 = 9125$$
$$\overline{X}^2 = 3025$$

$$S = \sqrt{\frac{9125}{3} - 3025} = \sqrt{16.7} = 4.1$$

Ambas distribuciones muestran la misma dispersión respecto a su media, es decir, tienen igual desviación estándar, sin embargo en términos relativos, una distribución donde el dato menor es 50 y el mayor 60 es mucho más homogénea (entendiéndose por homogénea que los datos bajos, intermedios y altos de la distribución son menos dispares entres sí) que otra donde el menor es 10 y el mayor 20. En el primer caso la diferencia entre el mayor y el menor es de 20% y en el segundo de 100%.

Esto revela que el calcular la desviación estándar es insuficiente para dar cuenta objetivamente de la dispersión de dos o más conjuntos de datos que se comparan entre sí. En consecuencia es necesario disponer de un indicador que tome en cuenta la tendencia central de la distribución. De aquí nace la necesidad de un **coeficiente de variabilidad.**

Coeficiente de variabilidad: Es la razón de la desviación estándar a la media de una distribución dada.
Se le conoce también como coeficiente de variación o desviación estándar relativa.
El coeficiente de variabilidad permite dar conclusiones más objetivas y se acostumbra expresarlo en %. Puede tomar valores muy grandes ya que no hay una relación de dependencia entre la S y la $\overline{X}$.

Si se obtiene el coeficiente de variabilidad de las dos distribuciones anteriores:

$$CV = \frac{S}{\overline{X}} = \frac{4.1}{15} = .273 = 27.3\%$$

$$CV = \frac{S}{\overline{X}} = \frac{4.1}{55} = 0.74 = 7.4\%$$

Dados al menos dos coeficientes de variabilidad, el menor de ellos pertenecerá a la distribución más homogénea, en este caso es la distribución 2.

CAPÍTULO V

PROBABILIDAD

5.1 INTRODUCCIÓN.

Cuando se presentó la definición de Estadística en el capitulo 2, se habló de las dos ramas en las que se divide ésta disciplina y que la Estadística inferencial descansa en la teoría de la probabilidad. Esta teoría tiene además muchas aplicaciones en diversos campos del quehacer humano.

La probabilidad como concepto intuitivo ha sido manejado por el hombre desde la más remota antigüedad, confundiéndola con ideas como: fortuna suerte, casualidad, etc. El primer trabajo de que se tiene memoria en el Mundo Occidental sobre el tema de lo probable, remonta al año 1654, cuando Pascal le escribió a Fermat, acerca de un problema, que le había planteado el caballero de Meré , y que consistía en repartir con justicia las apuestas de un juego de azar, interrumpido, y dando el puntaje a cada uno de los dos jugadores.

Fermat y Pascal se interesaron después en la teoría para determinar como repartir las apuestas de modo de" **estar seguro**" de ganar, jugando suficientemente en tiempo; es decir a la larga, cuando y cuanto está seguro de ganar.

5.1.1 FENÓMENOS DETERMINISTAS Y ALEATORIOS.

Existen respecto a la predicción dos clases de fenómenos. Por ejemplo si arrojamos un cerillo a un recipiente con gasolina, es inevitable que se incendie, si arrojamos una piedra al vacío sabemos de antemano que caerá, podemos incluso predecir a donde.

Estos experimentos cuyos resultados pueden ser anticipados con toda certeza reciben el nombre de:

- **Fenómenos deterministas.**

Sin en cambio si tiramos un dado en cuyas caras tenga pintado del 1 al 6, desconocemos cual de ellos quedará hacia arriba, o si echamos un volado, no estamos seguros de que saldrá águila o sol.
Estos experimentos en los que no es posible adelantar el resultado con toda certidumbre, se llaman:

- **Fenómenos aleatorios.**

Y estos son el objeto de estudio de la teoría de Probabilidad.

5.2 DEFINICIÓN Y PROPIEDADES DE LA PROBABILIDAD.

Si se tiene un evento que puede ocurrir de n maneras todas igualmente posibles, llamando s al número de veces que ocurre y f al número de veces que no ocurre, tendremos:

$$s+f=n$$

$$P(s)=\frac{s}{s+f}=\frac{s}{n}$$

$$P(f)=\frac{f}{s+f}=\frac{f}{n}$$

Que en palabras dice:

a) La suma de eventos n es igual a los que ocurren más los que no ocurren.
b) La probabilidad de que ocurra un evento es igual a los ocurridos entre los posibles.
c) La probabilidad de que no ocurra un evento es igual a los no ocurridos entre los posibles.

Resumiendo tendremos que:

$$\textbf{PROBABILIDAD} = \frac{\textbf{POSIBILIDADES ESPECÍFICAS}}{\textbf{POSIBILIDADES TOTALES}}$$

Una probabilidad no puede ser cualquier número, digamos -5 ó 27%, sino un número real P, que se asigna a un evento A que tiene las propiedades siguientes:

1.- La Probabilidad de que ocurra A no puede ser menor que cero ni mayor que uno. **El cero indica la imposibilidad de ocurrencia del suceso; el uno la certidumbre de que ocurrirá.**

$$0 \leq P(A) \leq 1$$

2.- Para dos eventos A_1 y A_2 mutuamente excluyentes, a, la probabilidad de ocurrencia de uno u otro es igual a la suma de sus probabilidades separadas.

Es decir:

$P(A_1 \text{ o } A_2)=P(A_1)+P(A_2)$, si A_1 y A_2 son mutuamente excluyentes

La expresión mutuamente excluyentes quiere decir que si ocurre uno de los eventos de un espacio muestral, ninguno de los otros puede ocurrir al mismo tiempo; la ocurrencia de cualquiera de ellos excluye automáticamente la de los restantes, dicho de otro modo: La probabilidad de que sucedan A_1 y A_2 es cero.

Por otro lado si A_1 y A_2 son mutuamente excluyentes y exhaustivos, o sea, que juntos abarcan todo el espacio muestral, la probabilidad de ocurrencia de uno u otro sigue siendo la suma de sus probabilidades separadas, pero en este caso esa suma es igual a uno, es decir:

$P(A_1 \text{ o } A_2) = P(A_1)+P(A_2)= 1$, si A_1 y A_2 son mutuamente excluyentes y exhaustivos.

5.3 ESPACIO MUESTRAL.

Para que se tenga una definición más científica de Probabilidad es necesario definir los conceptos con que se trabaja, relacionando la teoría de conjuntos con **la teoría de Probabilidad.**

Espacio muestral: Es el conjunto universal; conjunto de todos los n elementos, sucesos, relacionados pero diferentes.

Punto muestra: Cada uno de los elementos del espacio muestral.

Experimento : Un experimento es un proceso por medio del cual se obtiene una observación (o medición). Toda acción bien definida por ejemplo un volado.

Suceso : Es una colección especifica de puntos muestrales. Cada uno de los resultados de un experimento.

Evento : Subconjunto formado por la selección de los elementos especificados.

En un experimento, no siempre las observaciones tienen que ser numéricas por ejemplo:

1.- El puntaje obtenido en un examen.
2.- Lanzar al aire una moneda y observar si es cara o cruz.
3.- La inspección de bombillas para determinar si el producto es defectuoso o no.

Para poder contestar cualquier pregunta relacionada con el experimento, es primordial la correcta obtención del espacio muestral de este.

Comenzaremos por presentar experimentos sencillos y la obtención del espacio muestral relacionados a estos.

Ejemplo1:

Experimento: Se lanza al aire una moneda.

a)¿Cuáles son los posibles resultados?
b)¿Cuál es la probabilidad de obtener águila?

Solución :

Lo primero que haremos será encontrar el espacio muestral para poder contestar las dos preguntas anteriores.

Suceso	Moneda	Probabilidad
A_1	Sol	$P(A_1)=1/2$
A_2	Águila	$P(A_2)=1/2$
		$2/2=1$

Recuerda que para obtener la probabilidad de cada suceso solo aplicamos la definición de Probabilidad (Posibilidades especificas entre posibilidades totales). Para el primer suceso solo se tiene un sol y el total de posibilidades es dos. (1/2)
Ahora respondiendo a las dos preguntas

a)Los posibles resultados son dos: Águila y Sol.

b)La probabilidad de obtener águila es de ½

Ejemplo 2:

Experimento: Se lanza al aire un dado y se observa el número que aparece en la cara superior.

a) Haga una lista de los puntos muestrales.
b) ¿Cuál es la probabilidad de obtener un 6?
c) ¿Cuál es la probabilidad de obtener un número impar?

Solución:

Primero encontraremos el espacio muestral que en realidad es lo que se pide en el inciso a).

Suceso	Dado	Probabilidad
A_1	1	$P(A_1)=1/6$
A_2	2	$P(A_2)=1/6$
A_3	3	$P(A_3)=1/6$
A_4	4	$P(A_4)=1/6$
A_5	5	$P(A_5)=1/6$
A_6	6	$P(A_6)=1/6$
		6/6=1

b)La probabilidad de obtener un 6 es de 1/6 (es el sexto renglón del espacio muestral).

c)La probabilidad de obtener un número impar es de 3/6. Que son los renglones correspondientes a A_1 , A_3 y A_5.

A cada punto del espacio muestral le asignamos un número llamado la Probabilidad de A_i y denotado por el símbolo $P(A_i)$, tal que:

$$\sum P(A_i) = 1$$

Ejemplo 3:

Experimento: Se lanzan al aire tres monedas y se observa la cara superior.

a)¿Cuántos puntos muestrales son?
b)¿Qué probabilidad hay de que salgan tres soles?

Solución:

a)Para obtener el número de puntos muestrales, primero obtendremos el espacio muestral. Observando la tabla son 8 puntos muestrales.

Suceso	Dado	Probabilidad
A_1	A A A	$P(A_1)=1/8$
A_2	A A S	$P(A_2)=1/8$
A_3	A S A	$P(A_3)=1/8$
A_4	A S S	$P(A_4)=1/8$
A_5	S S S	$P(A_5)=1/8$ ←
A_6	S S A	$P(A_6)=1/8$
A_7	S A S	$P(A_7)=1/8$
A_8	S A A	$P(A_8)=1/8$
		8/8=1

b)La Probabilidad de que salgan tres soles es 1/8 (un suceso especifico entre 8 posibles).

Ejemplo 4:

Experimento: Se tiene en una urna, una baraja española (La baraja española consta de 4 palos diferentes. Se extrae una carta de la urna.

a)¿Cuál es la probabilidad de que la carta elegida sea un caballo?
b)¿Cuál es la probabilidad de que la carta sea de oros?
c)¿Cuál es la probabilidad de que no sea caballo?

Solución:

a)La probabilidad de sacar un caballo es de 4/40. Esto es porque solo existen 4 caballos (Uno de oros, uno de espadas, uno de bastos y otro de copas) de un total de 40 cartas.

b) La probabilidad de que la carta sea de oros es de 10/40, ya que existen 10 cartas del palo de oros entre un total de 40.

c) La probabilidad de que no sea caballo es de 36/40=9/10.

Ejemplo 5:

Experimento: Un inversionista tiene la oportunidad de invertir en tres de cinco acciones, el inversionista no sabe que solo dos de los cinco tipos producirán un beneficio sustancial en los primeros 5 años. Si el inversionista selecciona los tres tipos de acciones al azar. ¿Cuál es la probabilidad de que el inversionista escoja los dos tipos de acciones más beneficiosas?

Solución:

Obtendremos primero el espacio muestral del experimento.

Sean T_1 , T_2 , T_3 ,T_4 y T_5 las acciones y las acciones más beneficiosas T_1 y T_2.

S = Si es seleccionada

N = No es seleccionada.

Suceso	T_1	T_2	T_3	T_4	T_5	Probabilidad
A_1	S	S	N	N	N	$P(A_1)=1/10$
A_2	N	S	S	N	N	$P(A_2)=1/10$
A_3	N	N	S	S	N	$P(A_3)=1/10$
A_4	N	N	N	S	S	$P(A_4)=1/10$
A_5	S	N	S	N	N	$P(A_5)=1/10$
A_6	S	N	N	S	N	$P(A_6)=1/10$
A_7	S	N	N	N	S	$P(A_7)=1/10$
A_8	N	S	N	S	N	$P(A_8)=1/10$
A_9	N	S	N	N	S	$P(A_9)=1/10$
A_{10}	N	N	S	N	S	$P(A_{10})=1/10$

$10/10=1$

La Probabilidad de que el inversionista elija los dos tipos de acciones más beneficiosas es de 1/10.

Ejemplo 6:

Experimento : Cuatro trabajadores, de los cuales dos pertenecen a un grupo minoritario se asignan a cuatro empleos netamente diferentes.

a)Define el experimento.
b)Haga una lista de los puntos muetrales.
c)Si la asignación de empleos es imparcial, es decir si una asignación es tan probable como otra ¿Cuál es la probabilidad de que los dos empleados del grupo minoritario sean asignados a los dos empleos menos deseados.?

Solución:

a) El experimento es asignar 4 empleos a cuatro trabajadores.

b) En realidad están pidiendo el espacio muestral del experimento

Comenzaremos por definir los trabajos:

Sean E= Electricista, M = Mecánico , J = Jardinero y A = Afanador y los empleos menos deseados son J y A.

Y los trabajadores T_1 , T_2, T_3 y T_4 , siendo los del grupo minoritario T_1 y T_2, una vez definido lo anterior encontraremos el espacio muestral.

Suceso	A	E	M	J	Probabilidad
A_1	T_1	T_2	T_3	T_4	$P(A_1)=1/24$
A_2	T_1	T_2	T_4	T_3	$P(A_2)=1/24$
A_3	T_1	T_3	T_2	T_4	$P(A_3)=1/24$
A_4	T_1	T_3	T_4	T_2	$P(A_4)=1/24$
A_5	T_1	T_4	T_2	T_3	$P(A_5)=1/24$
A_6	T_1	T_4	T_3	T_2	$P(A_6)=1/24$
A_7	T_2	T_1	T_3	T_4	$P(A_7)=1/24$
A_8	T_2	T_1	T_4	T_3	$P(A_8)=1/24$
A_9	T_2	T_3	T_1	T_4	$P(A_9)=1/24$
A_{10}	T_2	T_3	T_4	T_1	$P(A_{10})=1/24$
A_{11}	T_2	T_4	T_1	T_3	$P(A_{11})=1/24$
A_{12}	T_2	T_4	T_3	T_1	$P(A_{12})=1/24$
A_{13}	T_3	T_1	T_2	T_4	$P(A_{13})=1/24$
A_{14}	T_3	T_1	T_4	T_2	$P(A_{14})=1/24$
A_{15}	T_3	T_2	T_1	T_4	$P(A_{15})=1/24$
A_{16}	T_3	T_2	T_4	T_1	$P(A_{16})=1/24$
A_{17}	T_3	T_4	T_1	T_2	$P(A_{17})=1/24$
A_{18}	T_3	T_4	T_2	T_1	$P(A_{18})=1/24$
A_{19}	T_4	T_1	T_2	T_3	$P(A_{19})=1/24$
A_{20}	T_4	T_1	T_3	T_2	$P(A_{20})=1/24$
A_{21}	T_4	T_2	T_1	T_3	$P(A_{21})=1/24$
A_{22}	T_4	T_2	T_3	T_1	$P(A_{22})=1/24$
A_{23}	T_4	T_3	T_1	T	$P(A_{23})=1/24$
A_{24}	T_4	T_3	T_2	T_1	$P(A_{24})=1/24$

c) La probabilidad de que los dos trabajos menos deseados sean asignados a los trabajadores del grupo minoritario es de 4/24 ó 1/6 es decir:

$$P(C) = 1/6$$

5.4 ANÁLISIS COMBINATORIO.

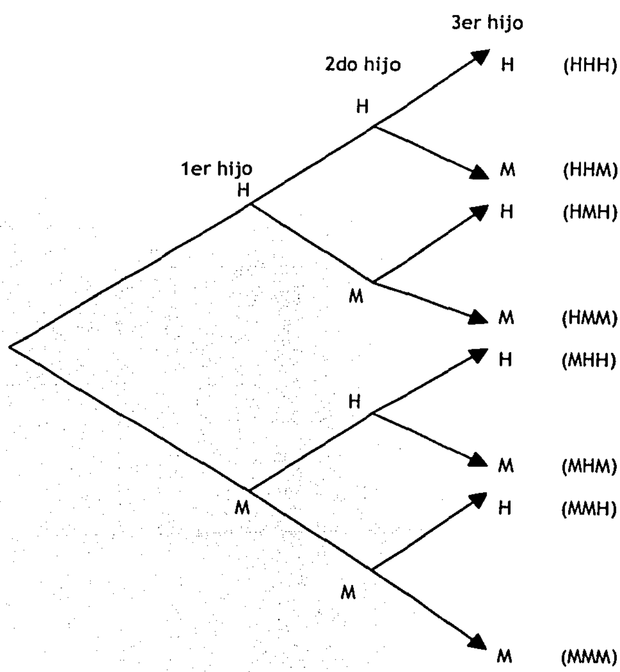
Al analizar la mayoría de los ejercicios anteriores, la obtención del espacio muestral no es muy complicado, pero en el último ejercicio se observa lo complejo que puede ser encontrar el espacio muestral para ciertos problemas. Por lo cual haremos uso del análisis combinatorio para la obtención de los espacios muestrales muy grandes, dicho de otra manera es una forma sofisticada de contar.

Se presentaran tres resultados matemáticos que son de gran utilidad para la resolución de problemas de probabilidad que implica la obtención del espacio muestral .Comenzaremos con diagrama de árbol.

5.4.1 DIAGRAMA DE ÁRBOL.

Si un evento A puede realizarse de n_1 maneras diferentes y una vez que esto ha ocurrido otro evento B puede realizarse de otras n_2 formas diferentes, entonces ambos eventos pueden realizarse en el orden indicado $n_1 \cdot n_2$ maneras distintas.
Este principio puede generalizarse para el caso que se tengan más de dos eventos.
Un diagrama de árbol es llamado así debido a su apariencia, se emplea frecuentemente para la resolución de ciertos problemas. A continuación presentaremos los siguientes ejemplos:

Ejemplo 1: Cierta pareja (o matrimonio), planea procrear tres hijos. ¿De cuántas maneras diferentes puede suceder este hecho?. Dicho de otra manera ¿Cuál es el espacio muestral?
Solución:
Sea H el evento de ser hombre y M de ser mujer, es evidente que el primer hijo puede ser tanto de un sexo como del otro, lo cual se ilustra con las primeras dos ramas, construyendo el diagrama de árbol:



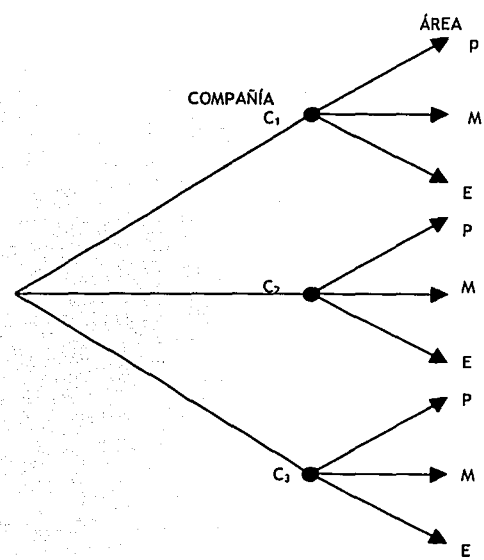
Siguiendo la ruta de cada rama podemos encontrar el espacio muestral:

$$n_1 \cdot n_2 \cdot n_3 = 2 \cdot 2 \cdot 2 = 8$$

Suceso	1°	2°	3°	Probabilidad
A_1	H	H	H	$P(A_1)=1/8$
A_2	H	H	M	$P(A_2)=1/8$
A_3	H	M	H	$P(A_3)=1/8$
A_4	H	M	M	$P(A_4)=1/8$
A_5	M	H	H	$P(A_5)=1/8$
A_6	M	H	M	$P(A_6)=1/8$
A_7	M	M	H	$P(A_7)=1/8$
A_8	M	M	M	$P(A_8)=1/8$

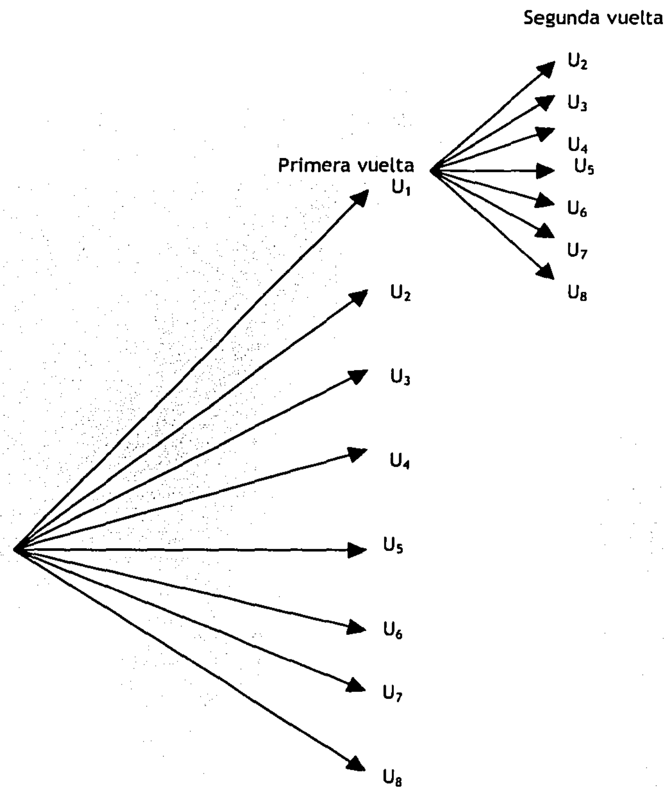
Ejemplo 2: 3 compañías tienen un empleo disponible en cada una de las siguientes áreas: Personal, mantenimiento y eléctrica ¿Cuántas oportunidades de empleo disponibles hay?

Solución: Calculando los valores $n_1 \cdot n_2 = 3 \cdot 3 = 9$ Entonces se tendrán 9 oportunidades de empleo.



Ejemplo 3: Una liga de futbol cuenta con 8 universidades participantes. Cada equipo debe jugar dos veces con cada universidad (Un juego de local y otro de visitante) ¿Cuántos partidos en total se jugaran en el torneo?

Solución: Podemos encontrar el espacio muestral por medio del uso de diagrama de árbol.



Al realizar el calculo $n_1 \cdot n_2 = 8 \cdot 7 = 56$

En total se jugarán 56 partidos en el torneo.

Antes de presentar el tema de permutaciones se explicará el proceso matemático del factorial.
En más de un proceso matemático aparecen expresiones matemáticas en las que hay que multiplicar los n primeros números enteros no negativos por ejemplo:

$1 \cdot 2 \cdot 3 \cdot 4 \ldots$ o bien $1 \cdot 2 \cdot 3 \ldots\ldots 9 \cdot 10$

Estas expresiones se pueden escribir como sigue:

4! , 10!

En general la expresión n! Indica: $n(n-1)(n-2)(n-3) \cdot \ldots\ldots\ldots \cdot 1$

Para poder generalizar algunas fórmulas, a 0! se le da el valor de 1.
El símbolo ! (factorial) crece muy rápidamente. La propiedad fundamental de n! es que:

$$n! = n(n-1)!$$

5.4.2 PERMUTACIONES

Otro de los resultados matemáticos que es de gran ayuda en la obtención de espacios muestrales y probabilidades, es la permutación.

Una permutación es un arreglo " ordenado" de r objetos distintos, este arreglo ordenado puede ser de una parte de los elementos, o de todos los elementos del conjunto.

Existen varias formas de representar una permutación, siendo la más difundida la siguiente:

$$_nP_r$$

en donde:
n = Número total de elementos del conjunto.
r = Número de elementos que se toman ala vez.
$_nP_r$= Número de permutaciones.

La fórmula anterior tiene la siguiente condición : $r \leq n$

Supongamos que se tienen n objetos, para formar cada permutación, podemos tomar n objetos para ocupar el primer lugar, hecho esto, el segundo lugar solo puede ser ocupado por (n-1)objetos, de modo que estos dos primeros lugares pueden tener n(n-1) posibilidades y para ocupar el tercer lugar tendremos (n-2) objetos, de modo que los tres primeros lugares tienen:
n(n-1)(n-2) posibilidades etc.

De modo que se tendrá $_nP_n$=n!

Que es el número de permutaciones de n objetos tomados de n en n.

Las permutaciones de n objetos tomados de r en n se obtiene de la siguiente formula:

$${}_{n}P_{r}=\frac{n!}{(n-r)!}$$

Aplicaremos las fórmulas anteriores a los siguientes ejemplos:

Ejemplo 1: Una persona tiene 4 libros V_1 , V_2 , V_3 y V_4.¿De cuántas formas se pueden colocar los libros tomando 2 cada vez ?

Solución: Se trata de una permutación ya que se van a ordenar los libros y como solo se va a ordenar una parte del conjunto se tiene que utilizar, la siguiente fórmula: ${}_{n}P_{r}=\frac{n!}{(n-r)!}$

Sustituyendo en la fórmula:

$${}_{4}P_{2}=\frac{n!}{(n-r)!}=\frac{4!}{(4-2)!}=\frac{4\cdot3\cdot2\cdot1}{2!}=\frac{24}{2\cdot1}=12$$

Presentando las 12 posibilidades se tienen:

V_1 V_2	V_2 V_3
V_2 V_1	V_3 V_2
V_1 V_3	V_2 V_4
V_3 V_1	V_4 V_2
V_1 V_4	V_3 V_4
V_4 V_1	V_4 V_3

Ejemplo 2: Un aparato está compuesto de 5 partes que pueden ser ensambladas en cualquier orden: se desea realizar una prueba para determinar el tiempo necesario para cada una de las formas posibles de montaje. Si cada forma se va a probar una sola vez ¿Cuántas pruebas deben realizarse?

Solución: El número total de pruebas se calcula con la fórmula ${}_{n}P_{n}=n!$, ya que todas las partes serán ensambladas.

$${}_{5}P_{5}= n!=5!=120$$

Serán 120 pruebas a realizarse.

Ejemplo 3: Un vendedor en México prepara un itinerario para visitar 6 ciudades principales. La distancia recorrida y por tanto el costo del viaje, dependerá de que ciudad se visita primero, segundo, ..., sexto lugar.¿Cuántos itinerarios distintos (costos distintos) son posibles?

Solución:

$$_6P_6=n!=6!=720$$

Se tienen 720 itinerarios distintos.

Ejemplo 4: Llegan a cierta ciudad 10 viajeros. ¿De cuantos modos podrán hospedarse en los 13 hoteles que hay, si solo ha de haber más que un viajero por hotel?

Solución:

$$_{13}P_{10}=\frac{13!}{(13-10)!}=103{,}783{,}600$$

5.4.3 COMBINACIONES.

Llamaremos combinación al número de subconjuntos de cardinalidad r que se pueden formar de un conjunto de cardinalidad n.

La fórmula que se utiliza para las combinaciones son las siguientes:

$$_nC_r=\frac{n!}{(n-r)!r!}$$

Como el orden no se considera en una combinación, es evidente que solo hay una combinación en que entren todos los objetos.

Aplicando la fórmula para las combinaciones, también encontramos que el número de ellas es uno.

$$_nC_n=\frac{n!}{(n-n)!n!}=\frac{1}{0!}=1$$

Recuerda que 0! = 1

Aplicaremos estas fórmulas a los siguientes ejemplos.

Ejemplo 1: Un entrenador de futbol tiene un equipo formado por once jugadores, entre los cuales está su hijo.

a) ¿Cuántas quintetas diferentes podrá formar con los 11 jugadores?
b)¿Cuántas quintetas diferentes podrá formar si su hijo siempre habrá de formar parte de la quinteta ?

Solución:

a) Se tienen 11 jugadores de los cuales hay que tomar 5 cada vez.

$_nC_r = \dfrac{n!}{(n-r)!r!}$ Sustituyendo los valores:

$_{11}C_5 = \dfrac{11!}{(11-5)!5!} = 462$ quintetas.

b) En este caso solo tendremos que calcular las combinaciones para 10 Jugadores, escogiendo 4, ya que el hijo siempre ha de formar parte de la quinteta.

$_{10}C_4 = \dfrac{10!}{(10-4)!4!} = 210$ quintetas.

Ejemplo 2: Se realizará un estudio de mercado. La población ha estudiar cuenta con 35 personas. Si se van a elegir muestras de 10 personas. ¿Cuántas muestras diferentes se pueden seleccionar?

Solución:

Calculando las combinaciones posibles de las 35 personas tomando 10 cada vez.

$_{35}C_{10} = \dfrac{35!}{(35-10)!10!} =$

Ejemplo 3: Con el propósito de determinar la habilidad de un testigo para identificar a 3 sospechosos de robo, se confronta al testigo con 10 hombres. supongamos que los 3 sospechosos que cometieron el crimen se encuentran entre los 10 hombres. Si el testigo en realidad no puede identificar a los sospechosos pero se siente obligado a hacer una elección.

a) ¿Cuál es la probabilidad de que los 3 hombres culpables sean seleccionados por azar ?
b) ¿Cuál es la probabilidad de que el testigo seleccione a 3 hombres inocentes?

Solución: Primero tenemos que calcular las posibilidades totales.

$$_{10}C_3=\frac{10!}{(10-3)!3!}=120$$

a) Solo existe una combinación en la que estén incluidos los tres sospechosos de robo. por lo que la probabilidad es.

$$P(a)=\frac{1}{120}$$

b)Para resolver este inciso es necesario primero calcular las combinaciones de las posibles combinaciones de las personas inocentes en el grupo:

$$_7C_3=\frac{7!}{(7-3)!3!}=35$$

Calculando la probabilidad:

$$P(b)=\frac{35}{120}=\frac{7}{24}$$

Ejemplo 4: Si en la zona geográfica por visitar hay 10 ciudades ¿Cuántas diferentes agrupaciones de 6 ciudades susceptibles de ser visitadas por el representante de ventas hay?

Solución :

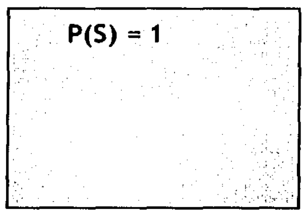
$$_{10}C_6=\frac{10!}{(10-6)!6!}=210$$

Existen 210 diferentes agrupaciones de ciudades para poder visitar.

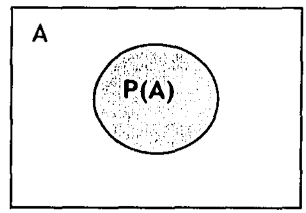
5.5 RELACIÓN ENTRE EVENTOS.

Existen varias relaciones que se dan entre eventos de un espacio muestral. A continuación se presentan por medio del uso de diagrama de Venn, ayudando a comprender mejor la representación gráfica de eventos y la relación entre ellos. El espacio muestral se simboliza por una S. Puesto que definir un espacio muestral es incluir todos los resultados posibles de un experimento, la probabilidad de que el resultado de cualquier intento dado provenga del espacio muestral es por fuerza igual a 1.

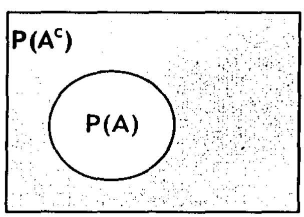
$$P(S) = 1$$



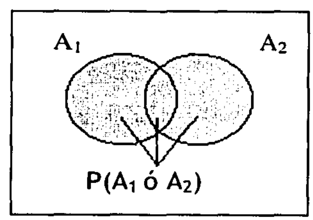
Espacio muestral



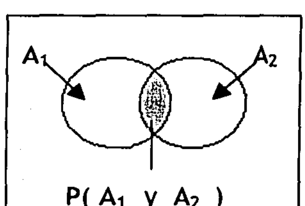
Evento A



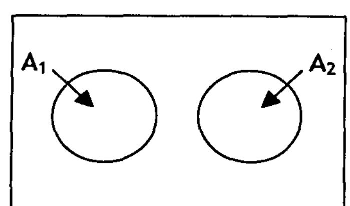
Complemento de A



Unión de A_1 y A_2



Intersección de A_1 y A_2



Eventos mutuamente excluyentes

El complemento de un suceso A es la colección de todos los puntos muestrales en S que no están en A. El complemento de A se denota con el símbolo $\overline{A}$. Esto se muestra a continuación.

$$\sum P(A_i) = 1$$
$$P(A) + P(\overline{A}) = 1$$
$$entoces$$
$$P(A) = 1 - P(\overline{A})$$

Esta relación es útil para obtener P(A) cuando se conoce $P(\overline{A})$ o viceversa.

Intersección:

Sean A y B dos sucesos en un espacio muestral S. La intersección de A y B es el suceso compuesto por todos los puntos muestrales que están en A y en B. La intersección de A y B se representa con el símbolo AB (Algunos autores usan), nosotros haremos uso de ambas.

Eventos independientes: Cuando la ocurrencia o no ocurrencia de un **suceso no tiene ningún efecto en la probabilidad de ocurrencia del otro suceso**, se dice que ambos eventos son independientes.
Por ejemplo si echamos un volado dos veces, es claro que si cae sol en el primero, esto no afecta la probabilidad de que en el segundo volado caiga águila. Para calcular la probabilidad de que ambos eventos sean independientes es:

$$P(A \cap B) = P(A) \cdot P(B)$$

Ejemplo 1: Si se tira un dado dos veces hallemos la probabilidad de que en ambas tiradas aparezca el uno.

Solución:

Es claro que la ocurrencia del uno no afecta la probabilidad de la ocurrencia del mismo número en la segunda tirada. Por lo tanto existe la independencia, por lo que calculando la probabilidad para estos eventos se tiene:

$$\text{P(1 y 1)} = \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{1}{36}$$

La probabilidad de que en ambas tiradas salga uno es de $\frac{1}{36}$

Eventos dependientes: Cuando **la ocurrencia o no ocurrencia de un suceso afecta a la probabilidad de ocurrencia del otro suceso** decimos que los eventos son dependientes. Por ejemplo si en el juego de domino mezclamos las fichas según las normas convencionales, la probabilidad de extraer cualquier ficha en el primer intento es de 1/28; si no se integra, la probabilidad de extraer algunas de las restantes en el segundo intento es de 1/27, pero ya que se tiene la primera ficha, la probabilidad de que esta misma salga en el segundo intento **es cero.**
Cuando dos eventos son dependientes, se emplea el concepto de probabilidad condicional, para designar la probabilidad de ocurrencia del suceso relacionado.

$$P(A \setminus B)$$

Donde la barra vertical dentro del paréntesis se lee " dado", y los sucesos que aparecen ala derecha de la barra son los que han ocurrido.entonces tendremos la siguiente definición:

$$P(B \setminus A) = \frac{P(AB)}{P(A)}$$

y

$$P(A \setminus B) = \frac{P(AB)}{P(B)}$$

Donde $P(AB) = P(A \cap B)$

Ejemplo 2: En el lanzamiento de dos dados Encontrar la probabilidad de que caigan dos números iguales con la condición de que la suma sea mayor que 9.

Solución :

Sea A el suceso de que caigan dos números iguales.

A= {(1,1);(2,2);(3,3);(4,4);(5,5);(6,6)}

Sea B el suceso de que la suma de los números sea mayor que 9.

B= {(4,6);(5,5);(6,4);(5,6);(6,5),(6,6)}

P(B) = 6/36 ; P(A)=6/36

$A \cap B = \{(5,5);(6,6)\}$

$P(A \cap B) = 2/36$

Ahora calcularemos $P(A \setminus B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{P(AB)}{P(B)}$

$$P(A \setminus B) = \frac{2}{6}$$

Aplicando la definición de probabilidad condicional tendremos la siguiente definición:

Decimos que dos sucesos A y B son independientes si:

P(A/B)=P(A)

ó

P(B/A)=P(B)

Si esto no ocurre decimos que los sucesos son dependientes.

Eventos mutuamente excluyentes: Dos o más eventos son mutuamente excluyentes, si no pueden ocurrir al mismo tiempo, esto es la ocurrencia de uno, excluye la ocurrencia de los demás.

Decimos que dos sucesos A y B son mutuamente excluyentes si el suceso AB no contiene ningún punto muestral.

La probabilidad del evento A o B, osea de que ambas puedan ocurrir, se expresa:

$$P(A \cup B) = P(A) + P(B) = 0$$

Ya que la intersección de ambas es un conjunto vacío o cero.

Ejemplo 3: El experimento consiste en lanzar un dado. Dados los siguientes sucesos sea:

A: Observar un número impar.
B: Observar un número menor que 4.

¿Los sucesos A y B son mutuamente excluyentes?, ¿ Son complementarios?, ¿Son independientes?

Solución:

La probabilidad del suceso A es de 1/2 y la probabilidad del suceso B es de 1/2 también. A continuación se da el espacio muestral para una mejor comprensión.

Espacio muestral:

SUCESO	DADO	PROBABILIDAD
E_1	1	1/6
E_2	2	1/6
E_3	3	1/6
E_4	4	1/6
E_5	5	1/6
E_6	6	1/6

O sea el suceso A está determinado por: E_1 , E_3 , y E_5, el suceso B está determinado por E_1 , E_2 y E_3 .

El suceso AB es el conjunto de puntos muestrales que están en A y B. Se observa que AB contiene los puntos E_1 y E_3. Por lo tanto A y B no son mutuamente excluyentes.
No son complementarios porque B no es conjunto de todos los puntos de s que no están en A.

De tal forma que haciendo los cálculos obtendremos lo siguiente:

$$P(A) = \frac{1}{2}$$

$$P(B) = \frac{1}{2}$$

$$P(AB) = P(A \cap B) = \frac{1}{3}$$

$$P(A \setminus B) = \frac{P(A \cap B)}{P(B)} = \frac{1/3}{1/2} = \frac{2}{3}$$

Para ver que sean independientes se tiene que cumplir que $P(A \setminus B) = P(A)$ lo cual no es cierto ya que $\frac{2}{3} \neq \frac{1}{2}$, por lo tanto como A y B dependen uno del otro no son independientes.

Ejemplo 4: Un supermercado tiene un remate de lastas sin etiqueta. Entre las latas que se van a rematar se encuentran 200 latas de maíz, 300 latas de remolacha y 500 latas de duraznos. ¿Cuál es la probabilidad de que la primera ama de casa seleccione una lata de vegetales?, ¿De maíz?, ¿Son estos dos sucesos independientes?, ¿Son mutuamente excluyentes?, suponiendo que el ama de casa selecciona una lata de vegetales, ¿Cuál es la probabilidad de que sea maíz?

Solución:

Sea A el suceso de que la primera lata que seleccione sea de vegetales.

P(A)= 500/1000=1/2

Sea B el suceso de que la primera lata que escoja sea de maíz.

P(B)= 200/1000=1/5

No son sucesos independientes puesto que A depende de B y no son mutuamente excluyentes puesto que la intersección si contiene puntos en común.

¿Cuál será la probabilidad de que sea una lata de maíz dado que seleccionó una de vegetales?

P(B\A)= $P(B \setminus A) = \dfrac{2/10}{1/2} = \dfrac{4}{10} = \dfrac{2}{5}$

De este mismo resultado podemos comprobar que no son independientes por medio de la siguiente igualdad:

$P(B \setminus A) = P(B)$ y con los resultados obtenidos $P(B \setminus A) = \dfrac{2}{5}$ y $P(B) = \dfrac{1}{5}$ se

confirma que no son iguales por lo tanto no son independientes.

Ejemplo 5: Una encuesta entre consumidores en una comunidad de los Estados Unidos, mostró que el 10% no quedó satisfecho con los trabajos de plomería hechos en sus casas. Cincuenta por ciento de las quejas fueron respecto al plomero A. Si el plomero A hace el 40% de los trabajos de plomería de la ciudad ¿Cuál es la probabilidad de obtener?

a) Un trabajo de plomería no satisfactorio, dado que el plomero es A.
b) Un trabajo satisfactorio, dado que el plomero es A.

Solución:

Sea A:Las quejas recibidas acerca del plomero A
Sea B:El suceso de obtener un trabajo no satisfactorio.

La probabilidad de obtener un trabajo no satisfactorio dado que es acerca del plomero A

Dicho de otra manera lo que se pide es: $P(B \setminus A) = \frac{P(A \cap B)}{P(A)} = \frac{.05}{.4} = .125$

Para obtener la probabilidad de un trabajo satisfactorio dado que fue hecho por el plomero A solo utilizamos P = 1-A^c

P(C) = 1 - 0.125 = .875

CAPÍTULO VI

DISTRIBUCIONES DE PROBABILIDAD

6.1 DISTRIBUCIÓN DE PROBABILIDAD.

La distribución de probabilidad de una variable **es una descripción de las probabilidades con que ocurren los diversos valores o categorías de esa variable.**

Las expresiones expuestas en la sección 5.8 son usadas tanto en las probabilidades discretas como en las probabilidades continuas, estas últimas son la que se utilizan en los trabajos estadísticos, en donde por lo general se requiere tener un modelo matemático que represente la distribución empírica de probabilidades.

Si consideramos el ejemplo siguiente de: Una pareja que desea procrear 3 hijos. El espacio muestral, está dado por 8 resultados igualmente probables: HHH , HHM , HMH ,HMM , MHH ; MHM; MMH y MMM , donde M es mujer y H es hombre. Considerando el número de hombres de estos 8 resultados posibles, 1 está formado por tres hombres; 3 por 2; 3 por 1 y 1 por 0. Por lo tanto aplicando, la interpretación clásica de probabilidad, conocemos que las:

1/3, 3/8, 3/8 y 1/8

Esto constituye la distribución de probabilidad que puede ser representada en forma de tabla como sigue:

DISTRIBUCIÓN DE PROBABILIDAD DEL NÚMERO DE HIJOS VARONES SI SE PROCREAN TRES

NUMERO DE VARONES	PROBABILIDAD
0	1/8 =.125
1	3/8 = .375
2	3/8 = .375
3	1/8 = .125
Total	**8/8 = 1**

En estadística aplicada, se conoce como variable aleatoria, cualquier variable cuyos valores puedan describirse mediante una distribución de probabilidad.

Partiendo de la base de sus posibles valores, las variables se clasifican en discretas o contínuas. Las variables discretas son las que se limitan a valores enteros, por otro lado las variables contínuas pueden tomar cualquier valor dentro de un intervalo.
En resumen, los valores de una variable discreta, son puntos desconectados en la recta numérica; pero los posibles valores de una variable contínua, incluyen todos los valores dentro de una región o intervalo ininterrumpido del eje real.

Un modo conveniente de presentar una distribución discreta de probabilidad es por medio de su función de probabilidad.

Una función de probabilidad: Es una función matemática que da la probabilidad de ocurrencia de cada valor de una variable aleatoria discreta.

Ejemplo: Considere el lanzamiento de un dado, la distribución del número de puntos asociados es X = 1,2,3,4,5,6, la función de probabilidad que describe esta distribución es:

$$P(x) = 1/6 \quad \text{para } x = 1,2,3,4,5,6$$

Esta es una manera simplificada de escribir las probabilidades:

$$P(1) = 1/6, \quad P(2) = 1/6, \ldots\ldots., P(6) = 1/6.$$

La variable aleatoria es el número de puntos en la cara superior y la función de probabilidad P(x) da la probabilidad de ocurrencia de cada uno de los 6 resultados posibles de esta variable discreta.

Como los valores de una función de probabilidad son "probabilidades", deben cumplir las siguientes propiedades:

$$0 \leq P(x) \leq 1 \quad \text{para todos los valores de x.}$$

$$\sum P(x) = 1$$

$$P(x_i \text{ o } x_j) = P(x_i) + P(x_j)$$

Una función de distribución: Es una función que muestra probabilidades acumuladas hasta determinados valores de la variable aleatoria.

La función de distribución se designa por $F(x_i)$, da la probabilidad de obtener un valor de la variable aleatoria x menor o igual a x_i:

$$F(x_i) = P(x \leq x_i)$$

Ya que los valores de la función de distribución $F(x_i)$ son probabilidad, nunca son menores de cero:

$$F(x) \geq 0$$

Cuando la variable x está en su valor máximo posible , x_{max}, el valor de$F(x)$, alcanza también su máximo de 1:

$$F(x_{max})=1$$

Ejemplo: Construyamos la función de distribución de todos los valores de la variable aleatoria " número de varones" al procrear tres hijos.

Solución : Sea x la variable aleatoria " número de varones ".

Entonces: x = 0,1,2,3

Sabemos que la probabilidad de tener 0,1,2, ó 3 hijos varones, es de 1/8,3/8,3/8 y 1/8 respectivamente. Quedando la función de distribución como sigue:

F(0)=P(0)=1/8
F(1)=P(0)+P(1)
F(1)=1/8+3/8=1/2

Dicho de otra manera la probabilidad de procrear un hijo varón o menos es de1/2.

F(2)=P(0)+P(1)+P(2)
F(2)=1/8+3/8+3/8=7/8

La probabilidad de procrear dos hijos varones o menos es de 7/8.

F(3)=P(0)+P(1)+P(2)+P(3)
F(3)=1/8+3/8+3/8+1/8=1

La probabilidad de procrear tres hijos varones es de 1 lo cual es lógico ya que F(3) es $F(x_{max})$.

Un concepto fundamental en la teoría de probabilidad es el de la esperanza matemática o valor esperado.

El valor esperado E(x) de una variable es el valor promedio de la variable que se obtendría a la larga.

Para calcular el valor esperado de una variable aleatoria x, se multiplica cada valor de la variable por su probabilidad de ocurrencia p sumándose luego todos los productos resultantes de los valores posibles de la variable.

$$E(x) = \sum x_i p_i$$
$$= x_1 p_1 + x_2 p_2 + \ldots\ldots + x_m p_m$$

para m valores mutuamente excluyentes y exhaustivos de x.

La media de una distribución de probabilidad se define como el valor esperado de la variable aleatoria. :

$$\mu = E(x)$$

La varianza de una distribución de probabilidad se define como el valor esperado (la media) de las desviaciones cuadráticas con respecto al valor medio:

$$\sigma^2 = E(x_i - X)^2$$
$$\sigma^2 = \sum (x_i - \overline{X}) p_i$$

La desviación estándar de una distribución de probabilidad es σ , la raíz cuadrada positiva de la varianza.

6.2 DISTRIBUCIÓN BINOMIAL.

Existen muchas distribuciones discretas de probabilidad, pero quizás, la más importante por sus numerosas aplicaciones en problemas empíricos, es la distribución binomial, desarrolla por Jacobo Bernoulli y publicada en 1713.

La distribución binomial es la distribución teórica de probabilidad del número de **éxitos** en una serie de repeticiones independientes, donde la probabilidad de éxito en cada repetición permanece constante.

La distribución binomial ocurre si estamos interesados en el número de veces que sucede un evento A en n ejecuciones independientes, de un experimento aleatorio.

Un experimento binomial, es un experimento que tiene las siguientes propiedades:

1.- El experimento consiste de n ensayos idénticos.
2.- Cada ensayo tiene dos resultados posibles, a uno de ellos lo llamaremos éxito y a otro fracaso.

3.- La probabilidad de éxito en un ensayo es igual a p y permanece igual en todos los ensayos. La probabilidad de fracaso es igual a q = 1-p.
4.- Los ensayos son independientes.

La forma general de la función de probabilidad binomial es:

$$f(x) = \frac{n!}{X!(n-x)} p^x \cdot q^{n-x}$$

Recordaras que la primera parte de la función de probabilidad binomial, corresponde a la fórmula de combinaciones visto en análisis combinatorio.

$$C_x^n = \frac{n!}{(n-x)!x!}$$

Recuerda que 0! =1

Ejemplo 1: Retomando el experimento de procrear tres hijos, y suponiendo que la probabilidad de engendrar un hijo varón es de 0.5, determinemos la probabilidad de que el número de varones sea:

a) 0
b) 1
c) 2
d) 3

Solución: Este experimento es aplicable a este tipo de distribución ya que consiste en tres repeticiones independientes en donde la probabilidad de éxito es engendrar 1 hijo varón es decir p = 0.5, por lo que q =1-p =1 - 0.5 = 0.5.

Obteniendo la probabilidad de que el número de varones sea 0 es:

$$P(0) = {}_3C_0 \cdot p^0 q^3$$

$$P(0) = 1\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^3$$

$$P(0) = 1(1)\left(\frac{1}{8}\right) = \frac{1}{8}$$

Calculando la probabilidad de que el número de varones sea 1:

$$P(1) = {}_3C_1 \cdot p^1 q^2$$

$$P(1) = 3\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^2 = \frac{3}{8}$$

Para el inciso c con 2 varones :

$$P(2) = {}_3C_2 \cdot p^2 q^1$$

$$P(2) = 3\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1 = \frac{3}{8}$$

Para el inciso d con 3 varones :

$$P(3) = {}_3C_3 \cdot p^3 q^0$$

$$P(3) = 1\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^0 = \frac{1}{8}$$

Es evidente que el utilizar la distribución binomial para obtener las probabilidades es más ventajosa y rápida que la manera anteriormente utilizada.

La media de una distribución binomial, designada por μ, es el valor esperado del número de éxitos .

$$\mu = np$$

La varianza de una distribución binomial es :

$$\sigma^2 = npq$$

La desviación estándar será la raíz cuadrada de la varianza:

$$\sigma = \sqrt{npq}$$

Ejemplo 2: El censo de 1990 el 90% de la población mexicana de 5 años de edad y más, profesa la religión católica y el resto otra o ninguna, si elegimos al azar 4 personas de 5 años de edad o más, determinemos la probabilidad de que entre los 4 sean católicos.

a) 0 ; b) 1 ; c) 2 ; d) 3 ; e) 4 .

Solución:

Este problema claramente se puede resolver por medio de una distribución binomial, ya que consiste de 4 intentos independientes, el éxito se tendrá cuando se profese la religión católica cuya probabilidad es p = 0.9, para cada intento es siempre la misma, por lo que el resultado desfavorable es q = 0.1

Para P = 0

$P(0) = {}_4C_0(.9)^0(.1)^4$

$P(0) = 1(1)(.0001) = .0001$

Para P = 1

$P(1) = {}_4C_1(.9)^1(.1)^3$

$P(1) = 4(.9)((.001) = .0036$

Para P = 2

$P(2) = {}_4C_2(.9)^2(.1)^2$

$P(2) = 6(.81)(.01) = .049$

Para P = 3

$P(3) = {}_4C_3(.9)^3(.1)^1$

$P(3) = 4(.729)(.1) = .291$

Para P = 4

$P(4) = {}_4C_4(.9)^4(.1)^0$

$P(4) = 1(.6561)(1) = .66$

La probabilidad de que entre las 4 personas de 5 años y más, elegidas al azar, haya exactamente 0 que practican la religión católica es de .0001; de que haya exactamente 1 es de .0036; que haya dos exactamente es .049, de que sean exactamente tres es .29 y por último para que sean exactamente cuatro es de .66.

La función de distribución correspondiente será:

F(0)=P(0) =.0001
F(1)=P(0)+P(1) =.0001+.0036
F(2)=P(0)+P(1)+P(2) =.0001+.0036+.049 =.053
F(3)=P(0)+P(1)+P(2)+P(3) =.0001+.0036+.049+.29 =.343
F(4)=P(0)+P(1)+P(2)+P(3)+P(4) =.0001+.0036+.049+.29+.66=1

La media de esta distribución se calcula de la siguiente manera:

$$\mu = np$$
$$\mu = 4(.9) = 3.6 \quad \text{católicos.}$$

Esto significa que a la larga se podría esperar una media de 3.6 católicos de cada 4 personas cuyas edades fuesen de 5 años y más .

Calculando la desviación estándar de este ejemplo :

$$\sigma = \sqrt{npq}$$
$$\sigma = \sqrt{4(.9)(.1)} = \sqrt{.36} = .6$$

En conclusión los datos se desvían .6 católicos en promedio del valor esperado.

6.2.1 ESTANDARIZACIÓN DE LA Z.

Hemos trabajado hasta ahora los conceptos de distribución, media, varianza y desviación estándar, conceptos indispensables y con los cuales el lector debe de estar ya familiarizado. El procedimiento conocido como estandarización forma parte de los conceptos que el lector debe manejar.

En estadística se entiende por transformación un conjunto de operaciones aritméticas que se realizan sobre los valores de una variable para obtener un nuevo conjunto de valores. A continuación se mostrará el procedimiento conocido como estandarización, que consiste en convertir datos como los que hasta ahora se han manejado en **datos estándar.**

Para transformar un conjunto de datos se hará lo siguiente:

1.- Se resta la $\overline{X}$ de cada dato.
2.- Se divide la diferencia de $X - \overline{X}$, entre la desviación estándar de la distribución.

Simbólicamente un dato estandarizado está dado por:

$$Z = \frac{X - \overline{X}}{S}$$

Si convertimos todas las categorías de una distribución en datos Z (otra manera de llamar a los datos estándar 9, obtendremos una nueva distribución en la cual la media y la desviación estándar valen siempre cero y uno respectivamente.

Es está característica o propiedad la que hace sumamente útiles a los datos Z. A continuación se dará la demostración matemática.

Dado un conjunto de datos Z, su media está dada por:

$$\overline{Z} = \frac{\sum Z}{N} \qquad (1)$$

Por otro lado un dato Z se define de la siguiente manera:

$$Z = \frac{X - \overline{X}}{S} \qquad (2)$$

Entonces podemos expresar la media de los datos Z en términos de los datos originales, solo sustituyéndolos la (2) en (1).

$$\overline{Z} = \frac{\sum \frac{X - \overline{X}}{S}}{N}$$

Ahora bien, como la desviación aparece como divisor de todas las diferencias $X - \overline{X}$, podemos extraerla como sigue:

$$\overline{Z} = \frac{\sum \frac{X - \overline{X}}{S}}{N} = \frac{\sum (X - \overline{X})}{NS}$$

Pero recordaras que al estudiar desviación estándar, obtuvimos que la sumatoria de todas las desviaciones con respecto a la media es siempre 0. Por lo tanto:

$$\overline{Z} = 0$$

De ésta propiedad resulta otra: Si la media de un conjunto de datos Z es cero, entonces su sumatoria por fuerza es cero.

$$\text{Si } \frac{\sum Z}{N} = \overline{Z} = 0, \text{ entonces } \sum \overline{Z} = 0$$

Ahora demostraremos que la desviación estándar de un conjunto Z es siempre 1; pero antes recordaremos que la varianza, por definición es la media de las desviaciones cuadráticas : $S^2 = \frac{\sum (X - \overline{X})^2}{N}$, de los cual se deduce que la sumatoria de las desviaciones cuadráticas es igual a la varianza por el número de datos.

$$\sum(X-\overline{X})^2 = NS^2$$

Y por definición podemos expresar la varianza de un conjunto de datos Z de la siguiente forma:

$$S_Z^2 = \frac{\sum(Z-\overline{Z})^2}{N}$$

Pero como la media de un conjunto de valores Z es cero, la expresión anterior se reduce a lo siguiente:

$$S_Z^2 = \frac{\sum Z^2}{N} \qquad (3)$$

Si expresamos la varianza de los datos Z en términos de los datos originales, sustituimos 2 en 3:

$$S_Z^2 = \frac{\sum Z^2}{N} = \frac{\sum\left(\frac{X-\overline{X}}{S}\right)^2}{N}$$

$$S_Z^2 = \frac{\sum\frac{(X-\overline{X})^2}{S^2}}{N} = \frac{\sum(X-\overline{X})^2}{NS^2}$$

Pero como recordemos que NS^2, solo es la sumatoria de las desviaciones cuadráticas, en consecuencia:

$$S_Z^2 = \frac{\sum(X-\overline{X})^2}{\sum(X-\overline{X})^2} = 1$$

De ésta propiedad se deduce otra también muy importante: Si la varianza y la desviación estándar de los datos Z son 1, entonces la sumatoria del cuadrado de los datos es igual al total de los datos.

Si $\frac{\sum Z^2}{N} = S_Z^2 = 1$, entonces $\sum Z^2 = N$

Resumiendo los resultados anteriores tenemos:

$$\overline{Z} = 0$$
$$\sum Z = 0$$
$$S_Z^2 = S_Z = 1$$
$$\sum Z^2 = N$$

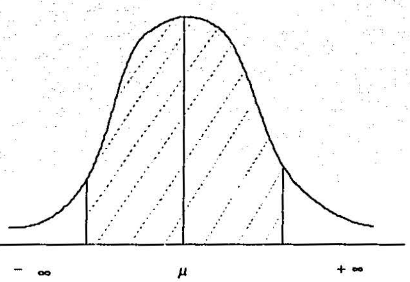
6.3 DISTRIBUCIÓN NORMAL.

Recordaremos que las variables aleatorias continuas están asociadas, con espacios muestrales que representan al número infinitamente grande de puntos muestrales contenidos en un intervalo de línea recta.
Un gran número de variables aleatorias que se observan en la naturaleza, poseen una distribución de frecuencias que es aproximadamente acampanada y continua.

Es decir recibe el nombre de distribución normal toda distribución acampanada y continua, debido a su continuidad se traza como una curva suave y no como un histograma .Su expresión matemática es la siguiente:

$$Y = \frac{e^{-\frac{(y-\mu)}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$$

Gráficamente la distribución normal es como se muestra:



La ecuación de la función de densidad está construida de manera que el área bajo la curva representa probabilidades; por lo tanto el área total es igual a 1. Aunque las distribuciones normales pueden tener medias y desviaciones estándar cualesquiera, es muy importante el caso donde la media y la desviación estándar

son 0 y 1 respectivamente. Esta distribución se conoce como **distribución normal estandarizada**, el área bajo la curva normal entre la media (z =0) y un valor especifico de z>0, digamos z_0, es la probabilidad.
Debido a su utilidad se han construido tablas que muestran el área bajo la curva limitada por dos ordenadas cualesquiera. esta tablas pueden ser usadas en todo conjunto de datos distribuidos normalmente, luego de haber sido estandarizados.

Estas tablas tienen las siguientes características:

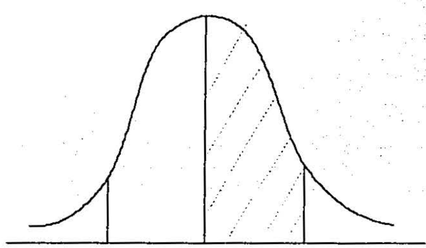
a) El área total bajo la curva es 1 ó 100% y, puesto que es simétrica el área a ambos lados de la media es 0.5 ó 50%.
b) Solo la mitad de la curva normal está representada en la tabla ya que es simétrica, no es necesaria la otra mitad, para encontrar algunas áreas será necesario solo hacer algunas operaciones aritméticas.

La tabla de distribución normal estandarizada aparece en el apéndice de este libro.
Para estandarizar utilizaremos:

$$z = \frac{y-\mu}{\sigma}$$

Ejemplo 1: Hallar el área bajo la curva limitada a un lado por la media y al otro por z = 1.5.

Solución : Gráficamente lo que se pide es lo siguiente :



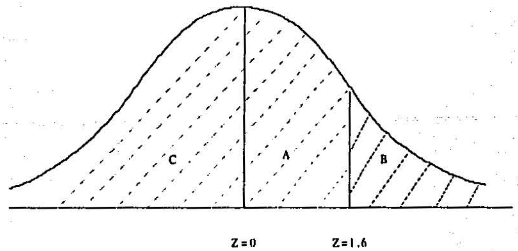
Buscando en la tabla de distribución normal en z =1.5 y la columna 00, encontramos el siguiente valor z = .4332 que es .4332 = 43.32%.

Esto significa que el 43.32 % de todos los datos de la distribución tienen valores comprendidos entre la media y 1.5 desviaciones estándar.

Ejemplo 2: Determinar el área que se haya

a) Por arriba de z = 1.6
b) Por debajo de z = 1.6

Representamos gráficamente lo que se pide:



Solución :

a) El área que tenemos que encontrar se halla por arriba de z = 1.6, o sea el área representada por B, encontrando el valor de z =1.6.

z = 1.6 es .4452, recordando que la mitad equivale a 0.5, solo se restará el valor encontrado a 0.5.

0.5 -.4452 = .0548

Esto significa que el 5.48 % de los datos de la distribución tienen valores no menores de 1.6 desviación estándar.

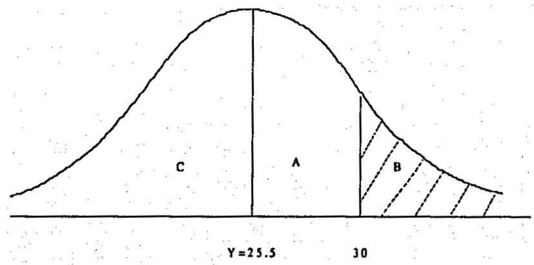
b) Primero obtendremos el valor para z = 1.6 , que es .4452 que es el 44.52%, el área que se nos pide es el resultado de sumar C+A, es decir:

0.5 + .4452 = .9452 = 94.52%

Lo que significa que el 94.52 % de los datos de la distribución tienen valores no mayor de 1.6 desviaciones estándar.

Ejemplo 3: Ciertos estudios demuestran que el consumo de gasolina de los autos medianos tiene una distribución normal con un consumo medio de 25.5 Km. por galón y una desviación estándar de 4.5 Km. por galón. ¿Qué porcentaje de autos medianos obtiene 30 ó más Km. por galón?

Solución: Gráficamente lo que se pide es lo siguiente. El área sombreada es lo que se está buscando, primero encontraremos el valor de z que corresponde a y = 30 Km./galón:



Sustituyendo en la fórmula para Z:

$$z = \frac{y-\mu}{\sigma} = \frac{30-25.5}{4.5} = 1.0$$

Una vez encontrado el valor de z, buscaremos este valor en las tabla de distribución normal .

Para z = 1 es .3413, pero como nosotros solo queremos el área sombreada recordemos que la mitad es igual a .5 por lo que:

0.5 -.3413 =.1587

Dicho de otra manera el 15.87 % excede a 30 Km./galón.

6.4 INTERVALOS DE CONFIANZA.

Los intervalos de confianza son límites probabilísticos que se pueden asignar a un pronostico o a una predicción con el fin de tener un marco de referencia sobre el cual tomar decisiones.
Al establecer tales limites de confianza es conveniente tener en cuenta que existe una relación inversa entre el nivel de confianza y la precisión del pronostico. **En efecto a mayor confianza menor precisión.**

Explicaremos esto con un ejemplo: Suponga que usted está jugando rayuela en el salón de clases ¿Qué tanta confianza tiene usted de que al lanzar la moneda caiga justamente sobre la raya? Obviamente se debe aclarar que influyen factores tales como la experiencia y la puntería del jugador.
Sin embargo dejando de lado éstos factores, también es obvio que si la raya estuviera más ancha (digamos 10cm.). Sería mayor la confianza que usted tendría en acertar. ¿No es cierto?

Y por supuesto que si la raya fuera de un metro de anchura usted estuviese completamente seguro de que las monedas lanzadas caerían sin duda alguna sobre la franja.

Pues esto es precisamente lo que ocurre con los límites de confianza en el caso de un pronostico. Al investigador le interesaría construir un rango o intervalos dentro del cual pueda estar seguro de que ésta el valor verdadero y esto se llama **intervalo de confianza.**

Cuanto mayor número de muestras individuales se tengan, más fuertemente se concentrarán las medias muestrales alrededor de la media del universo.

El 90% caerá dentro del intervalo de ± 1.64 el error típico respecto a la media muestral.

El 95% caerá dentro del intervalo ± 2 el error típico, respecto de la media muestral.

El 99% o sea prácticamente la certeza de que la media del universo está comprendida en el intervalo ± 3 veces el error típico respecto de la media muestral.

Construcción de intervalos de confianza.

La media de cierto universo es desconocida, pero la desviación estándar de ese universo es conocida.

Con una probabilidad del 95%, la media muestral está situada entre

$$\mu \text{ universo } -\frac{2\sigma}{\sqrt{n}} \text{ y } \mu \text{ universo } +\frac{2\sigma}{\sqrt{n}} \quad \therefore \mu \pm \frac{2\sigma}{\sqrt{n}}$$

Si no se conoce la desviación estándar del universo, hay que sustituirlo por un estimador como la desviación estándar de la muestra.

$$\hat{S}_{\overline{X}} = \frac{\hat{S}}{\sqrt{n}}$$

donde:

$\hat{S}_{\overline{X}}$ = Desviación estándar de la media muestral.
$\hat{S}$ = Desviación estándar de la muestra.
n = Número de observaciones de la muestra.

Ejemplo 1: Una muestra al azar de 400 amas de casa suministra la siguiente información, cerca de la cantidad de dinero gastada en productos alimenticios durante 6 meses.

Media de la muestra $\bar{x} = \$400$
Desviación estimada de la muestra $S = \$80$

Estimar la media del universo con un intervalo de confianza del 95%

Solución:

$$\mu = \bar{x} \pm (2\hat{S}) = 400 \pm \frac{2S}{\sqrt{n}}$$

$$\mu = 400 \pm 2\frac{(80)}{\sqrt{400}} = 400 \pm \frac{2(80)}{20}$$

$$\mu = 400 \pm 8$$

Con un intervalo de confianza del 95%, la media del universo está entre \$392.00 y \$408.00.

Ejemplo 2: La producción diaria de un producto elaborado en una planta química para n =50 tuvo una media y desviación estándar de:

Media $\bar{x} = 871$ toneladas.
Desviación estándar $S = 21$ toneladas.

Encuentre un intervalo de confianza del 90% para la media de la población.
Solución :

Utilizando los limites de confianza para el 90% se tendrá:

$$\mu = \bar{x} \pm (1.64S) = 871 \pm 1.64\frac{(21)}{\sqrt{50}}$$
$$\mu = 871 \pm 4.89$$

Por lo tanto estimamos que la producción diaria promedio, se encuentra en el intervalo de 866.11 a 875.89 toneladas. El coeficiente de confianza .90 indica que al repetirse el muetreo, el 90% de los intervalos de confianza contendrá a μ.

CAPÍTULO VII

APLICACIONES DE PROBABILIDAD Y ESTADÍSTICA A LOS PROYECTOS DE INVERSIÓN.

En el capitulo uno hablamos acerca de la determinación del riesgo mediante el empleo de algunos conceptos estadísticos, tales como la media ponderada, la desviación estándar, la estandarización de valores, así como la búsqueda de valores bajo la curva normal tema visto en el capitulo seis. Todos los elementos de estadística y probabilidad pueden ser utilizados en el proceso de aceptación de un proyecto de inversión, desde el estudio de mercado hasta el tipo de riesgo que pueda presentarse al aceptar ciertos proyectos de inversión.
En este capitulo haremos uso de algunos elementos de estadística y probabilidad aplicados a un proyecto de inversión.

7.1 DETERMINACIÓN DEL RIESGO DE UN PROYECTO DE INVERSIÓN.

A continuación daremos la explicación de un ejemplo de la determinación del riesgo por medio de distribuciones de probabilidad, en donde aplicaremos algunos elementos de estadística y probabilidad.

1.- Primeramente defina el indicador financiero que habrá de utilizar para medir el riesgo: TIR, VAN, B/C, punto de equilibrio, una razón financiera etc. Por supuesto siempre debe preferir indicadores del más alto nivel, tales como la rentabilidad. Por ejemplo en este caso la Tasa interna de rendimiento **(TIR).**

TIR, Tasa interna de rendimiento es la tasa de descuento máxima que puede aceptar el proyecto.
VPN, Valor presente neto, es el beneficio del proyecto a valor presente ya descontado el costo de la inversión inicial.
TREMA, Tasa de rendimiento mínima atractiva

2.- Defina por lo menos tres escenarios alternos asociados con el entorno en que se desarrolla el evento o problema a resolver.
Los escenarios típicos son: pésimo, esperado y óptimo. Pero se pueden emplear otros tales como: Recesión, normal y auge. Por ejemplo:

ESCENARIO
OPTIMISTA O AUGE
ESPERADO O NORMAL
PESIMISTA O RECESION

3.- Defina la probabilidad de ocurrencia del evento, para cada uno de los escenarios considerados. Es obvio que la definición de los diversos escenarios requiere del concurso y de la experiencia del analista. De ésta forma la tabla quedaría de la siguiente manera:

ESCENARIO	PROBABILIDAD DE OCURRENCIA
OPTIMISTA O AUGE	25%
ESPERADO O NORMAL	50%
PESIMISTA O RECESION	25%

4.- Calcule el valor del indicador financiero de referencia para cada uno de los diferentes escenarios. En este caso, el indicador financiero será el rendimiento, mismo que puede estar expresado en términos absolutos o relativos, por ejemplo: VAN, relación de B/C, o TIR, tal como se ilustra a continuación:

ESCENARIO	PROBABILIDAD DE OCURRENCIA	TIR
OPTIMISTA O AUGE	25%	17.0
ESPERADO O NORMAL	50%	5.0
PESIMISTA O RECESION	25%	-7.0

5.- Calcule el valor esperado del indicador financiero empleado. Este valor esperado es simplemente un promedio ponderado del indicador financiero anteriormente definido es decir (la TIR promedio).

El factor empleado en la ponderación será el porcentaje de probabilidad definido para cada escenario y el resultado esperado se obtendrá en dos etapas, la primera se refiere a multiplicar la probabilidad de ocurrencia de cada evento por el indicador financiero obtenido en cada escenario en este caso la TIR; la segunda etapa se refiere a sumar los productos obtenidos. En términos generales el valor esperado promedio (k barra) se obtiene aplicando la siguiente expresión:

$$\bar{k} = \sum (k_X P_X)$$

donde:

k_X =Es el valor esperado del indicador financiero para el escenario X.
P_X = La probabilidad de ocurrencia del escenario X.

De manera especifica, para el caso que nos ocupa, la fórmula para calcular el valor esperado de la TIR quedaría en los siguientes términos:

$$TIR = \sum (TIR_X)(P_X)$$

Con esta información el ejemplo quedaría de la siguiente manera:

ESCENARIOS ALTERNOS	PROBABILIDAD DE OCURRENCIA P_x	RENDIMIENTO TIR_x %	VALOR ESPERADO DE CADA ESCENARIO O (p_x TIR_x)
OPTIMISTA	0.25	17.0	4.25
ESPERADO	0.50	5.0	2.50
PESIMISTA	0.25	-7.0	-1.75
TASA INTERNA DE RENDIMIENTO ESPERADA O PROMEDIO = 5.00			

El valor esperado de la rentabilidad (TIR promedio) representa el rendimiento promedio de todo el proyecto o problema en cuestión y el resultado se obtiene sumando algebraicamente los datos contenidos en la última columna, que para este caso resulta ser de 5.00

6.- Determina la dispersión o variabilidad absoluta y relativa del valor esperado promedio (k barra), mediante su varianza, su desviación estándar y su coeficiente de variación. Una mayor desviación estándar supone una mayor volatilidad de los flujos de efectivo esperados y en consecuencia un mayor riesgo. En este caso la varianza es la suma de la última columna esto es: 72.0 y dado que la desviación estándar es la raíz cuadrada de la varianza, entonces la desviación estándar es igual a 8.485281, como se ilustra a continuación:

ESCENARIOS ALTERNOS	P_x	TIR_x %	(P_x TIR_x)		(P_x TIR_x-TIR)^2
OPTIMISTA	**0.25**	**17.0**	**-1.75**	**0.25**	**$(17.0 - 5.0)^2 = 36.0$**
ESPERADO	**0.50**	**5.0**	**2.50**	**0.50**	**$(5.0 - 5.0)^2 = 0$**
PESIMISTA	**0.25**	**-7.0**	**4.25**	**0.25**	**$(-7.0 - 5.0)^2 = 36.0$**
					VARIANZA = 72.0
					DESV. ESTANDAR = 8.4853

El coeficiente de variación es el cociente del valor esperado promedio y la desviación estándar. Dado que este indicador se muestra en porcentaje, el resultado de la división se multiplica por 100.

$$CV = \frac{TIR}{S} \times 100$$

El indicador que mide más apropiadamente el riesgo es el coeficiente de variación. Entre la varianza y la desviación estándar hay que decidirse por la última debido a que la varianza expresa los resultados elevados al cuadrado lo cual no tiene una explicación económica o financiera. Sin embargo considere que

no es posible el cálculo de la desviación estándar sin haber calculado previamente la varianza. Por otra parte mientras la desviación estándar es un indicador absoluto, el coeficiente de variación es de tipo relativo.
Continuemos con el procedimiento.

7.- Normalice la variabilidad o dispersión del valor esperado promedio de acuerdo con el indicador financiero empleado y los criterios que éste establece para su aceptación. Dado que el indicador empleado es en este caso el TIR, la normalización se hará mediante la siguiente expresión algebraica:

$$Z = \frac{TREMA - TIR}{S_{TIR}}$$

donde:

TREMA = Tasa de rendimiento mínima atractiva.
S_{TIR} = Desviación estándar de la TIR esperada o promedio.

Así la normalización del rendimiento esperado se hará sustituyendo en la fórmula los valores de la TREMA, de la TIR promedio y de la desviación estándar de la TIR promedio. La TREMA se refiere al rendimiento mínimo que un inversionista desea para decidirse a invertir y que está determinado por el costo de oportunidad de sus recursos. Dado que estamos ilustrando un procedimiento asumiremos que para el caso que nos ocupa la TREMA es de 1.2% real.
Sustituyendo los valores numéricos en la fórmula el resultado sería el siguiente:

Z=-0.44783

Donde el signo simplemente indica que el valor encontrado se encuentra a la izquierda del centro de la curva normal. Si el resultado tuviese signo positivo significaría que en este caso tal valor se encuentra a la derecha del centro de dicha curva normal. El valor de Z debe leerse o entenderse como " el número de veces la desviación estándar"; en efecto el valor de -0.44783 significa que la diferencia de la TREMA y la TIR promedio (TREMA-TIR) es igual a -0.44783 desviaciones estándar. Lo anterior puede comprobarse usando los valores numéricos de las variables de la siguiente manera:

$$Z = \frac{1.2 - 5.0}{8.485281}$$

Multiplicando por Z el denominador (desviación estándar) tendremos que el primer miembro de la igualdad equivale a la diferencia de la TREMA y la TIR, que es precisamente lo que se indica en el párrafo anterior.

-0.44786Z=1.2-5.0 esto es:
-0.44783(8.485281)=1.2-5.0
-3.8=-3.8

Para el caso que nos ocupa, dado que el indicador financiero empleado fue la TIR, el indicador de referencia tuvo que ser la TREMA, pero si hubiese empleado otro indicador financiero el punto de referencia sería el parámetro que sirve de criterio para aceptar proyectos.

7.2 ACEPTACIÓN DE PROYECTOS DE INVERSIÓN.

Existe una gama de procedimientos que el administrador financiero puede aplicar para adecuarse a situaciones de riesgo e incertidumbre al realizarse el análisis y la evaluación de los proyectos de inversión, mencionados en el capitulo uno.

Uno de ellos es el método subjetivo, que se basa en la apreciación personal del especialista en la evaluación de proyectos, es decir considera al sujeto y no al objeto.

7.2.1 MÉTODO SUBJETIVO EN BASE A PROBABILIDADES.

Se puede aplicar el método subjetivo en base a dos criterios:

a) Análisis de sensibilidad.
b) Probabilidades. La probabilidad de que un evento ocurra puede ser considerada como un porcentaje de oportunidad de obtener un cierto resultado.

Partiendo de esta aseveración, el procedimiento consiste en considerar tres posibles alternativas de flujos netos de efectivo, posteriormente se cataloga a estos como pesimista, más probable y optimista, se otorga a cada uno de ellos un determinado porcentaje de probabilidad que en su conjunto deberá se igual al 100%; se multiplica cada flujo neto de efectivo por el porcentaje que le fue asignado, los resultados de las operaciones se suman obteniéndose así " el valor esperado más probable", es decir un promedio de las estimaciones de los flujos anuales, a esta cantidad se le aplica el factor de valor presente, originándose así el valor presente esperado más probable, por último a esta cantidad se le resta el monto de la inversión original, si la diferencia (valor presente neto) es positiva, el proyecto es aceptado de lo contrario se rechaza.

Ejemplo :

La Compañía HALLEY, S.A. de C.V., cuyo giro es la fabricación de enseres menores, desea adquirir una maquinaría computarizada. Se le presenta a el encargado de evaluar los proyectos de alternativas y se le indica que deberá analizarlos de acuerdo a su experiencia profesional seleccionando uno de ellos.
La tasa de descuento que deberá aplicar de acuerdo a las especificaciones que proporciona la entidad es de 95%.

Los datos proporcionados son:

- Monto de las inversiones correspondientes a los proyectos **X** y **Y**.
- Los flujos netos de efectivo generados por cada uno durante su vida útil.
- Tasas anuales de rendimiento pesimista, más probable y optimista.

Solución:

La siguiente tabla indica que de acuerdo al análisis de sensibilidad los proyectos **X** y **Y** tienen una amplitud de variación del 30% y 20% respectivamente, en base a esto se considera que el proyecto Y es menos riesgoso que el X.

Al aplicar probabilidades a los flujos netos esperados conforme a los criterios pesimista, más probable y optimista, el proyecto X tiene un valor presente neto de $ 231,886,000 y el proyecto Y de $ 22,336,000 o sea que si se llevase a cabo el proyecto X la ganancia que se tendría con respecto al Y sería de $ 209,520,000 de acuerdo a lo anterior el X sería el seleccionado.
Apoyándose en la aplicación del método subjetivo y relacionando la amplitud de variación de ambos proyectos así como el valor presente neto esperado más probable, el proyecto aceptado sería el X, puesto que la variación no se considera sumamente significativa mientras que la diferencia en los valores presentes netos más probables es bastante considerable. La siguiente tabla muestra lo anterior.

CIFRAS EN MILES DE PESOS

	P R O Y E C T O S	
	X	Y
INVERSIÓN ORIGINAL =	$ 800,000	$ 800,000
AÑO 1	$ 600,000	$ 1,100,000
AÑO 2	$ 700,000	$ 1,000,000
AÑO 3	$ 800,000	$ 900,000
AÑO 4	$ 900,000	$ 800,000
AÑO 5	$ 1,000,000	$ 700,000
AÑO 6	$ 1,100,000	$ 600,000
AÑO 7	$ 1,200,000	$ 500,000
AÑO 8	$ 1,300,000	$ 400,000
SUMAS	$ 7,600,000	$ 6,000,000
PROMEDIO DE LOS FLUJOS	$ 950,000	$ 750,000
TASA INTERNA DE RENDIMIENTO	88.00 %	127.67 %

ANÁLISIS DE SENSIBILIDAD

TASA ANUAL DE RENDIMIENTO:		
PESIMISTA	73.00 %	117.67 %
MÁS PROBABLE	88.00 %	127.67 %
OPTIMISTA	103.00 %	137.67 %
AMPLITUD DE VARIACIÓN	30.00 %	20.00 %

PROBABILIDAD

		FLUJO NETO DE EFECTIVO	**FLUJO NETO DE EFECTIVO**
PESIMISTA	20 %	$ 600,000	$ 400,000
MÁS PROBABLE	50 %	$ 950,000	$ 750,000
OPTIMISTA	30 %	$1,300,000	$1,100,000
		FLUJO MÁS PROBABLE	**FLUJO MAS PROBABLE**
PESIMISTA		$ 120,000	$ 80,000
MÁS PROBABLE		$ 475,000	$ 375,000
OPTIMISTA		$ 390,000	$ 330,000
VALOR ESPERADO MÁS PROBABLE		$ 985,000	$ 785,000
VALOR PRESENTE AL 95 % DEL VALOR ESPERADO MÁS PROBABLE		$1,031,886	$ 822,366
VALOR PRESENTE NETO		$ 231,886	$ 22,366

SE ACEPTA EL PROYECTO X

7.2.2 MÉTODO ESTADÍSTICO.

Otro de los procedimientos que el administrador financiero aplica para adecuar a situaciones de riesgo e incertidumbre al realizar el análisis y la evaluación de los proyectos de inversión es el método estadístico, que consiste en determinar la desviación estándar y el coeficiente de variación de cualquiera de los siguientes datos:

a) Flujos netos de efectivo.
b) Flujos netos de efectivo a valor presente.
c) Tasas internas de rendimiento.

Recordemos que la desviación estándar es el resultado de obtener la raíz cuadrada del promedio de los cuadrados de las desviaciones de cada uno de los valores con respecto a la media.

Entre mayor sea la desviación estándar y el coeficiente de variación, mayor será el riesgo del proyecto en cuestión y viceversa. La construcción de una tabla de distribución de frecuencias facilita más el trabajo para la determinación de los valores.

Ejemplo:

Tomaremos los datos de los proyectos del ejemplo anterior. Los pasos a seguir para resolver el problema mediante este método son los siguientes:

1.- En la primera columna se tienen los conceptos, en este casa los años durante los cuales se espera la generación de flujo de efectivo.

2.- En la segunda columna (Xi) los flujos de efectivo por cada año.

3.- En la tercera columna $(Xi-\overline{X})$ las cifras se calculan restando al flujo de efectivo de cada año el promedio de los flujos.

4.- En la cuarta columna $(Xi-\overline{X})^2$ se eleva al cuadrado el resultado obtenido de la tercera columna.

Solución :

Una vez realizada la tabla de distribución de frecuencias, se calculan la desviación estándar así como el coeficiente de variación, para llegar a una conclusión.

CIFRAS EN MILES DE PESOS

		PROYECTO X	
PERIODOS	Xi	$Xi-\overline{X}$	$(Xi-\overline{X})^2$
AÑO 1	\$ 600,000	-(\$ 350,000)	\$ 122,500,000,000
AÑO 2	\$ 700,000	-(\$ 250,000)	\$ 62,500,000,000
AÑO 3	\$ 800,000	-(\$ 150,000)	\$ 22,500,000,000
AÑO 4	\$ 900,000	-(\$ 50,000)	\$ 2,500,000,000
AÑO 5	\$ 1,000,000	\$ 50,000	\$ 2,500,000,000
AÑO 6	\$ 1,100,000	\$ 150,000	\$ 22,500,000,000
AÑO 7	\$ 1,200,000	\$ 250,000	\$ 62,500,000,000
AÑO 8	\$ 1,300,000	\$ 350,000	\$ 122,500,000,000
SUMAS	\$ 7,600,000	...\$ 0	\$ 420,000,000,000

PROMEDIO DE LOS FLUJOS $= \frac{\$7,600,000}{8}$ = 950,000

DESVIACIÓN ESTANDAR $= S = \sqrt{\frac{(Xi-\overline{X})^2}{n}} = \sqrt{\frac{420,000,000,000}{8}} =$ = 229,129

COEFICIENTE DE VARIACIÓN $CV = \frac{S}{\overline{X}} = \frac{229,129}{950,000}$ = 0.2412

PERIODOS	Xi	PROYECTO Y $Xi-\overline{X}$	$(Xi-\overline{X})^2$
AÑO 1	$ 1,100,000	$ 350,000	$ 122,500,000,000
AÑO 2	$ 1,000,000	$ 250,000	$ 62,500,000,000
AÑO 3	$ 900,000	$ 150,000	$ 22,500,000,000
AÑO 4	$ 800,000	$ 50,000	$ 2,500,000,000
AÑO 5	$ 700,000	-($ 50,000)	$ 2,500,000,000
AÑO 6	$ 600,000	-($ 150,000)	$ 22,500,000,000
AÑO 7	$ 500,000	-($ 250,000)	$ 62,560,000,000
AÑO 8	$ 400,000	-($ 350,000)	$ 122,500,000,000
SUMAS	$ 6,000,000	...$ 0	$ 420,000,000,000

PROMEDIO DE LOS FLUJOS $= \frac{\$6{,}000{,}000}{8}$ = 750,000

DESVIACIÓN ESTANDAR =S= $\sqrt{\frac{(Xi-\overline{X})^2}{n}} = \sqrt{\frac{420{,}000{,}000{,}000}{8}} =$ = 229,129

COEFICIENTE DE VARIACIÓN CV= $\frac{S}{\overline{\overline{X}}} = \frac{229{,}129}{750{,}000}$ = 0.3055

Se obtuvieron desviaciones estándar de 229,129.000 para ambos proyectos, debido que la diferencia entre los flujos de efectivo para cada año es la misma tanto en X como en Y, independientemente de los montos de dichos flujos, por lo que bajo este criterio los dos proyectos tienen la misma oportunidad de ser aceptados. En lo concerniente a los coeficientes de variación estos fueron de 24.12% y 30.55% para los proyectos X y Y respectivamente, de donde se concluye que el primero es menos riesgoso que el segundo en un 6.43%.

Una vez ponderados los dos criterios de aceptación en los que se basa el método estadístico, se determina que el proyecto **X será el seleccionado**.

Es importante destacar que estos dos parámetros toman en cuenta únicamente la distribución de los valores de la muestra o población, es decir la distancia que separa a los datos de su promedio y no consideran la cantidad y calidad del numerario.

Aunque estos ejemplos no son en lo único en que pueden ser aplicados los conocimientos de Probabilidad y Estadística sirven para darnos cuenta la utilidad de ellas en un proyecto de inversión.

CONCLUSIONES

Aprender a invertir y vigilar las inversiones requiere de un gran esfuerzo. Desde el hecho de tomar la decisión en que voy a invertir hasta el manejo de los resultados de la inversión, involucra el uso de las Matemáticas.

Es menos riesgoso emprender ésta labor, teniendo las bases matemáticas necesarias que nos permitan la realización de ésta tarea de una manera más eficaz, utilizando en este caso la aplicación de la **Estadística y la Probabilidad**, para obtener los mejores resultados posibles.

Estableciendo puntos de partida y la manera de analizar las distintas alternativas de inversión, podemos aplicar estás dos ramas de la Matemática en un proyecto de inversión, por ejemplo:

- La aceptación de un proyecto de inversión.
- El riesgo de un proyecto de inversión.
- Análisis de mercado.
- Control de calidad, etc.

Este trabajo da las herramientas básicas para poder aplicar La Estadística y Probabilidad a cualquier problema que así lo requiera, pudiéndose entre otras cosas aplicar a problemas de producción, de ventas, de investigación, de control de calidad etc., áreas íntimamente ligadas a un proyecto de inversión.

Por otro lado uno de los objetivos de este trabajo, ha sido desarrollar ejemplos que nos permitan involucrarnos en las áreas de Estadística y probabilidad de una manera sencilla y objetiva; los ejemplos se resolvieron de tal manera que al lector no se le dificulte la interpretación y resolución del problema, ya que han sido presentados en su mayoría paso a paso. Abarcando gran diversidad de las áreas en donde se pueden aplicar estas dos ramas de las matemáticas.

Aunque es importante recordar la gran ventaja que tiene el contar con los conocimientos básicos de Estadística y Probabilidad ya que al aplicarlos a la vida cotidiana, en el trabajo o escuela resulta de gran utilidad.

Recordemos que el Actuario tiene la capacidad de emprender ésta labor de iniciación en estas dos ramas, ya que su formación académica está íntimamente ligada a ellas.

Por otro lado en cuanto al desarrollo de temas financieros no existe ningún inconveniente para poderlos emplear, ya que está capacitado para hacerlo;

siendo la combinación de bases matemáticas con las financieras uno de los temas desarrollados en este trabajo, presentados en el último capítulo.

Así el Actuario pretende ayudar a iniciar y reforzar los conocimientos matemáticos mediante el desarrollo de este trabajo.

Ejercicio 2.4

1.- De los siguientes ejemplos anota en el paréntesis, la letra D si se trata de una variable aleatoria discreta o una C si es una variable aleatoria continua.

a)- Velocidad de un automóvil. ( )
b)- Cantidad de agua que se consume en la ciudad de Monterrey. ( )
c)- Tiempo de vuelo de un avión. ( )
d)- Número de libros consultados en un mes en una biblioteca. ( )
e)- Cantidad de alumnos egresados del CETIS 6 en el año de 1999. ( )
f)- Suma de puntos en el lanzamiento de un dado. ( )
g)- El número de bacterias por centímetro cúbico de agua potable. ( )
h)- La presión arterial. ( )
i)- La cantidad de lluvia que cae en la Ciudad de México durante una Semana. ( )
j)-La cantidad de alumnos de primer ingreso que se reciben cada año .( )

Ejercicios 2.6.2

2.- Para cada uno de los casos siguientes obtenga una distribución de frecuencias, obteniendo la frecuencia absoluta, frecuencia relativa y la frecuencia absoluta acumulada.

a) Calificaciones de Matemáticas, obtenidos por los alumnos de 3er semestre.

5, 5, 9, 8, 7, 7, 7, 6, 6, 6, 6, 8, 9, 10, 9
5, 5, 6, 6, 6, 7, 7, 8, 10, 7, 7, 8, 7, 7, 7

b) Horas de clase frente a grupo de los profesores de bachillerato.

4, 18, 12, 20, 5, 12, 15, 6
20, 24, 10, 11, 8, 16, 16, 5
12, 11, 10, 12, 8, 12, 13, 7

Ejercicio 2.6.4

3.- Obtenga una distribución de frecuencias por intervalos de los siguientes datos:

a) Calificaciones obtenidas al finalizar el bachillerato por los alumnos del CETIS #5.

Construya la distribución de frecuencias con 6 intervalos.

4.5	4.7	8.2	4.9	7.8	5.6	7.2	9.3	9.9	8.4	8.4	8.5
8.5	7.5	7.6	7.6	8.7	8.6	5.5	5.3	6.4	6.8	7.2	7.3
6.7	9.1	6.4	6.9	5.9	6.7	8.9	9.1	5.8	7.8	7.8	7.8
8.9	8.9	8.9	6.4	6.4	6.4	6.4					

b) Los datos siguientes representan el tiempo dedicado al estudio fuera de clase, en horas semanarias, por estudiantes universitarios construye la distribución de frecuencias con 7 intervalos.

3, 2, 5, 8, 2, 5, 1, 21, 7, 1, 11, 4, 3, 15, 4, 5, 16
13, 10, 8, 9, 20, 4, 3, 12, 1, 12, 23, 11, 22, 6, 17, 5, 6
2, 13, 8, 1, 10, 3, 7, 4, 2, 15, 6, 4, 14, 5, 12, 10, 5
2, 10, 17, 9, 2, 1, 6, 16, 1, 3, 18, 18, 3, 6, 1, 6, 11
4, 12

Ejercicio 2.6.5

4.- De la siguiente distribución de frecuencias que representa las edades de los trabajadores de la fabrica de tableros de control. Obtenga:

a) El histograma de frecuencia absoluta contra intervalo.

b) El polígono de frecuencia relativa contra intervalo.

INTERVALO	FRECUENCIA ABSOLUTA	FRECUENCIA RELATIVA
26 - 30	3	.142
30 - 34	5	.238
34 - 38	7	.333
38 - 42	3	.142
42 - 46	2	.095
46 - 50	1	.047

5.- La siguiente distribución de frecuencias contiene los datos de las personas que cometieron suicidio en el Estado de Colima.

INTERVALO	FRECUENCIA ABSOLUTA
16 - 20	12
20 - 24	11
24 - 28	5
28 - 32	2
32 - 36	2
36 - 40	3

Obtenga :
a) El histograma de: frecuencia relativa contra intervalos.

b) El polígono de frecuencias de: frecuencia contra intervalos.

Ejercicio 3.2

6.- Dada la siguiente sucesión encuentra la fórmula.

0,3,8,15,24.....

Ejercicio 3.3

7.- Obtenga las siguientes sumas:

a) $\sum_{y=0}^{5}(y-4)=$

b) $\sum_{x=2}^{6}(x^2-5)=$

Ejercicios 3.4.1

8.- Obtenga la media aritmética para los siguientes conjuntos de datos:

a) 5, 6, 8, 2, 3, 2, 1, 2, 9,3,7, 8,7

b) 300,345,675,234,5678,1234,1023,678,254,998,234,768,945,235,866

9.- De la siguiente distribución de frecuencias correspondientes al número de miembros con los que cuenta cada familia en cierto pueblo. Obtenga la media.

NUMERO DE MIEMBROS VARIABLE	FRECUENCIA
1	1
3	2
4	8
5	19
6	17
7	10
8	10
9	3
10	4
11	1
12	4

10.- De la siguiente tabla de distribución de frecuencias por intervalos correspondiente al número de hermanos que tienen los alumnos de preparatoria. Obtenga la media.

NUMERO DE HEMNAOS INTERVALO	FRECUENCIA	MARCA DE CLASE
0 - 2	31	1
2 - 4	72	3
4 - 6	41	5
6 - 8	12	7
8 - 10	2	9
10 - 12	2	11

Ejercicio 3.4.2

11.- Calcula la mediana para la siguiente tabla de distribución correspondiente a las jornadas de trabajo de los obreros de cierta fabrica de ropa.

HORAS DE TRABAJO	FRECUENCIA
4	5
6	22
8	18
10	15
12	12

12.-De la siguiente tabla de distribución de frecuencias por intervalos correspondientes alas horas semanarias dedicadas al estudio por parte de los alumnos universitarios. Obténgase la mediana.

HORAS SEMANARIAS INTERVALO	FRECUENCIA
1 - 3	50
4 - 6	38
7 - 9	26
10 - 12	36
13 - 15	19
16 - 18	7
19 - 21	6
22 - 24	5

Ejercicio 3.4.3

13.- De la siguiente tabla de distribución de frecuencias obtenga la moda. Los datos corresponden al salario mensual de los obreros de cierta compañía de lácteos.

SALARIO MENSUAL	FRECUENCIA
1400	7
1600	5
1800	9
2000	13
2200	8
2400	11
2600	7

14.-De la siguiente distribución de frecuencias por intervalos encuentra la moda. Los datos corresponden a la edad de alumnos en los grupos de primer año a nivel bachillerato técnico..

AÑOS DEL ALUMNO INTERVALO	FRECUENCIA
18 - 20	49
21 - 23	57
24 - 26	20
27 - 29	12
30 - 35	14

Ejercicio 4.2

15.- De la siguiente distribución de frecuencias por intervalos correspondiente a la antigüedad (o tiempo trabajado) en cierta institución de gobierno. Obtenga la desviación media.

ANTIGÜEDAD INTERVALO	FRECUENCIA
1 - 3	17
4 - 7	21
8 - 11	26
12 - 15	25
16 - 19	16
20 - 25	20

Ejercicio 4.3

16.-De la siguiente tabla de distribución de frecuencias por intervalos correspondiente al número de materias no aprobadas por los alumnos de bachillerato durante los tres años. Obtenga la Varianza.

MATERIAS NO APROBADAS INTERVALO	FRECUENCIA
0 - 3	13
4 - 7	21
8 - 11	19
12 - 15	16
16 - 19	17

Ejercicio 5.2

17.-De los siguientes números ¿Cuáles no pueden ser probabilidades ?

1.6 , 0.45 , 78% , 125.1% , 0 , 16/3 , -.025 , -4

Ejercicio 5.3

18.- ¿Qué probabilidad se tiene en el juego de domino al sacar una ficha ?

a) De que la ficha sea la de seis, tres.
b) De sacar la mula de seis.
c) De sacar cualquier mula.

19.- De acuerdo a estudios realizados a niños de 3 a 8 años, la posibilidad de ser derecho o zurdo es de 13 a 1 de que sea derecho. ¿Cuál es la probabilidad de ser zurdo?.

20.- En una urna se cuenta con una baraja española. Si se extrae una carta al azar. ¿Cuál es la probabilidad de que la carta ?

a) Sea un rey.
b) No sea un caballo.
c) Sea un basto.

21.- Se eligen a 5 estudiantes para formar una pareja y representar al colegio en un concurso de ciencias básicas. El grupo de los 5 estudiantes está formado por 3 hombres y 2 mujeres. ¿Cuál es la probabilidad de que la pareja elegida este formada por?

a) Un hombre y una mujer.
b) Dos mujeres.
c) Dos hombres.

22.- Una investigación realizada entre comerciantes ambulantes menores de 16 años revela que la probabilidad de que el vendedor no vaya a la escuela es de 3/16. Hallar la probabilidad de que si vaya.

23.- Se echa un volado 4 veces. Halle la probabilidad de obtener:
a) 0 águilas.
b) 1 águila o más.
c) 3 águilas o más.
d) 3 águilas y un sol en ese orden.
e) 2 águilas y 2 soles en cualquier orden.

Ejercicio 5.4.1

24.- Una pareja planea procrear dos hijos. ¿Cuál es la probabilidad de que sean las dos mujeres?. Resuélvalo por medio de diagrama de árbol.

25.- Una agencia de viajes, cuenta con tres rutas diferentes para ir de México a Boston, cinco rutas diferentes para ir de Boston a California y cuatro rutas distintas para ir de California a Toronto. ¿ Cuántas rutas diferentes e tiene en total si tienes que pasar de Boston a California y de California a Toronto?

26.- La compañía Valencia desea ascender a cuatro de sus gerentes para ocupar los siguientes puestos: Vicepresidente de Fianzas, Vicepresidente de cobranzas, Vicepresidente de Recursos humanos y Vicepresidente de ventas. ¿De cuántas maneras puede realizar estos ascensos?.

27.- Salen 8 viajeros en 8 barcos de Acapulco a Mazatlán ¿De cuántos modos podrán viajar, ida y vuelta, si un mismo viajero no debe ocupar el mismo barco que otro viajero, y no tomar el mismo barco de regreso, que él que tomó de ida?

28.- ¿Cuántas señales diferentes puedes realizar con 5 astas y 6 banderas, si han de utilizarse las 5 astas cada vez?

Ejercicio 5.4.3

29.- Se sabe que dos puntos determinan una recta ¿Cuántas rectas se pueden trazar con 6 puntos si nunca tres puntos están alineados?

30.- Se va a realizar un estudio para determinar la actitud de los empleados de cierta fabrica, respecto a los procedimientos administrativos a seguir durante las incapacidades medicas. Si se va a seleccionar una muestra de 7 empleados de un total de 88.¿Cuántas muestras distintas se pueden seleccionar?

31.- Con el propósito de determinar la habilidad de un testigo para identificar a tres sospechosos de robo, se confronta al testigo con 10 hombres. Supongamos que los tres sospechosos que cometieron el crimen se encuentran entre los 10 hombres. Si el testigo en realidad no puede identificar a los sospechosos pero se siente obligado a hacer una elección.
a) ¿Cuál es la probabilidad de que los tres hombres culpables sean seleccionados por azar?
b)¿Cuál es la probabilidad de que el testigo seleccione a tres hombres inocentes?

Ejercicio 5.5

32.- Un experimento genera un espacio muestral que contiene 8 sucesos E_1 ....E_8 con $P(E_i)=1/8$, i =1, ,8. Los sucesos A y B se definen como sigue:

A : E_1, E_4, E_6
B : E_3, E_4, E_5, E_6, E_7

Encuentre :
$a) P(A)$
$b) P(\overline{A})$
$c) P(A \cup B)$
$d) P(AB)$
$e) P(A \setminus B)$
f) ¿Son los sucesos A y B mutuamente excluyentes? ¿Por qué?
g) ¿Son los sucesos A y B independientes?¿Por qué?.

33.- Una encuesta entre consumidores en una población particular mostró que el 10 % no queda satisfecho con los trabajos de plomería hechos en sus casas. Cincuenta por ciento de las quejas fueron respecto al plomero A. Si el plomero A hace el 40 % de los trabajos de plomería de la ciudad, ¿Cuál es la probabilidad de obtener:

a) Un trabajo de plomería no satisfactorio, dado que el plomero es A?
b) Un trabajo satisfactorio dado que el plomero es A.?

Ejercicio 6.2

34.-Una encuesta en una ciudad particular mostró que 9 de 10 automóviles tienen seguro de responsabilidad civil. Si 4 autos en esta ciudad se ven involucrados en un accidente, ¿Cuál es la probabilidad de que?

a) No más de dos de los cuatro tengan seguro?
b) Exactamente dos tengan seguro?.

35.-Consideremos un defecto metabólico que ocurre en aproximadamente uno de cada 100 nacimientos. Si cuatro niños nacen en un hospital un día dado, ¿cuál es la probabilidad de que:

a) Ninguno tenga el defecto?
b) No más de uno tenga el defecto?

36.- Muchas compañías de electricidad han empezado a promover la conservación de energía ofreciendo tarifas reducidas a los consumidores que mantienen su consumo de energía debajo de ciertos límites. En un informe reciente, se observa que el 70 % de los residentes de Puerto Rico han reducido su consumo de electricidad lo suficiente para beneficiarse de las tarifas reducidas. Si se seleccionan 5 residencias al azar en San Juan, Puerto Rico encuentre la probabilidad de que:

a) Las 5 se beneficien de la tarifa reducida?.
b) Al menos cuatro se beneficien de la tarifa reducida?.

37.- Supongamos que el 80% de una raza de cerdos está infectado con triquinosis. Si se examina una muestra aleatoria de 1,000 cerdos:

a) ¿Cuál es el valor esperado del número y de cerdos infectados en la muestra?
b) ¿Cuál es la desviación estándar de y?
c) Si el inspector encuentra y = 900 cerdos infectados en la muestra, ¿podríamos creer que el porcentaje de cerdos infectados en la población es realmente 0%?

Ejercicio 6.3

38.- Las puntuaciones en una prueba nacional de aprovechamiento tuvieron una distribución normal con media de 500 y desviación estándar de 100. Si Pedro obtuvo 650 puntos ¿Qué fracción del total de estudiantes consiguieron una puntuación mayor que la de Pedro ?

39.- Determina el área bajo la curva normal delimitada como se indica:
a) Entre la media y z = 2
b) Entre la media y z = -1.8
c) Por debajo de z = 1
d) Por arriba de z = -2.1

40.- En una prueba de duración para determinar la vida útil de un nuevo tipo de lámpara lanzada al mercado, los datos de la muestra se distribuyeron de manera casi normal. Se calculo una vida útil promedio de 7,000 horas y una desviación de 520 horas. Halla el porcentaje de lámparas que tendrán una vida útil de:

a) Entre la media y 7,600 horas.
b) Entre 6,500 horas y 7,600 horas.
c) Más de 7,600.

Ejercicio 2.4

1.-
a) C
b) C
c) C
d) D
e) D
f) D
g) D
h) D
i) C
j) D

Ejercicio 2.6.2

2.-
a)

VARIABLE	FRECUENCIA	FRECUENCIA RELATIVA	FRECUENCIA ABS. ACUM.
5	4	.1333	4
6	7	.2333	11
7	10	.3333	21
8	4	.1333	25
9	3	.1000	28
10	2	. 0666	30
	30	.99	

b)

VARIABLE	FRECUENCIA	FRECUENCIA RELATIVA	FRECUENCIA ABS. ACUM.
4	1	.0416	1
5	2	.0833	3
6	1	.0416	4
7	1	.0416	5
8	2	.0833	7
10	2	.0833	9
11	2	.0833	11
12	5	.2083	16
13	1	.0416	17
15	1	.0416	18
16	2	.0833	20
18	1	.0416	21
20	2	.0833	23
24	1	.0416	24
	24	.999	

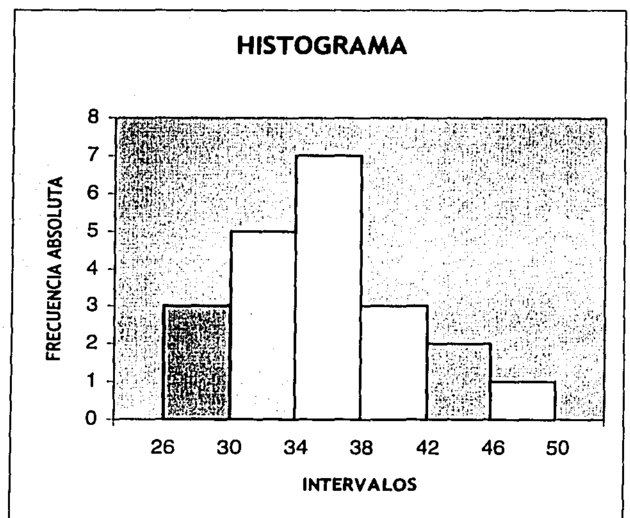
Ejercicio 2.6.4

3.-

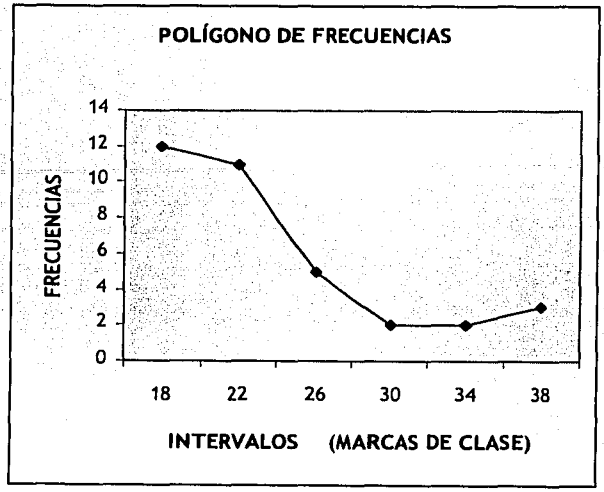
INTEVALO	FRECUENCIA ABSOLUTA	FRECUENCIA RELATIVA	FRECUENCIA ACUMULADA	FRECUENCIA REL. ACUMULADA.
4.5 - 5.5	5	.1162	5	.1162
5.5 - 6.5	9	.2093	14	.3255
6.5 - 7.5	8	.1860	22	.5115
7.5 - 8.5	11	.2558	33	.7673
8.5 - 9.5	9	.2093	42	.9766
9.5 - 10.5	1	.0232	43	.9998
	43			.999

Ejercicio 2.6.5

4.- a)



5.- b)



Ejercicio 3.2

6.- La fórmula es $f(y) = y^2 - 1$

7.- a) -9, b) 65

Ejercicio 3.4.1

8.- a) $\overline{X} = \frac{63}{13} = 4.84$

b) $\overline{X} = \frac{1447}{15} = 964.46$

9.- $\overline{X} = 56.60$

10.- $\overline{X} = 3.6$

Ejercicio 3.4.2

TESIS CON FALLA DE ORIGEN

11.- Me = 8

12.- Me = 7.42

Ejercicio 3.4.3

13.- Moda = $ 2,000

14.- Moda = 21.35

Ejercicio 5.2

1.6, 78 %, 125.1%, 16/3, -.025

Ejercicio 5.3

18.-
a) P(A) = 1/27
b) P(B) = 1/27
c) P(c) = 6/27

19.- P(A) = 13/14

20.-
a) P(A) = 1/10
b) P(B) = 9/10
c) P(C) = ¼

21.-
b) P(B) = 2/10
c) P(C) = 3/10

22.- P(A) = 13/16

23.-
a) P(A) = 1/16
b) P(B) = 15/16
c) P(C) = 5/16
d) P(D) = 7/16

Ejercicio 5.4.1

24.- P(A) = ¼

25.- 60 rutas diferentes.

26.- 24 formas de ascender a los trabajadores.

27.- 203,212,800

28.-720

Ejercicio 5.4.3

29.- 15

30.- 6348337336

31.-
a) P(A) = 1/120
b) P(B) = 7/24

Ejercicio 5.5

32.-
a) 3/8
b) 5/8
c) 3/4
d) 1/4
e) 2/5
f) no, E_4 y E_6 están en A y B
g)
$P(A \setminus B) \neq P(A)$
$P(B \setminus A) \neq P(B); no$

33.-
a) .125
b) .875

Ejercicio 6.2

34.-
a) .0523
b) .0486

35.-
a) .9606
b) .9941

36.-
a) .1681
b) .5282
37.-
a) 800
b) 12.649
c) No, y = 900está a cerca de 8 desviaciones estándar sobre la media.

Ejercicio 6.3

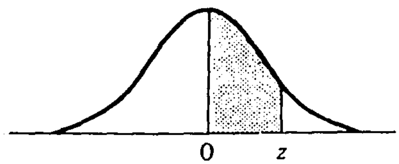
38.- .0668

39.-
a) .4773
c) .8413
d) .9821

40.-
a) 36.43%
b) 69.58%
c) 13.57%

APÉNDICE

Áreas bajo la curva normal :



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

BIBLIOGRAFÍA

GALLARDO, Cervantes Juan; Formulación y evaluación de proyectos de inversión. Mcgraw Hill. México 1998. pags. 70, 71, 72, 73, 86, 87, 88, 89, 90.

HURTA, Ríos Ernestina; Análisis y Evaluación de proyectos de inversión para bienes de capital. Instituto mexicano de contadores públicos. México. 1990. pags. 11, 12, 13, 15, 16, 111, 113, 114, 115.

LOHR, l. Sharon; Muestreo diseño y análisis. International Thomson editores. Estados Unidos. 2000. pags. 35, 36.

KREYSZIG, Erwin; Introducción a la Estadística matemática. Editorial Limusa S.A. de C.V. México. 1987. pags. 83, 84 , 97, 98, 193, 194.

MENDENHALL, William; Introducción a la Probabilidad y Estadística. Wadsworth International. Estados Unidos. 1982. pags.12, 13, 14, 15, 16, 30, 31, 32, 83, 84, 125, 126, 127, 158, 159, 196, 197, 198.

MAGAÑA, Cuellar Luis; Matemáticas III Estadística y Probabilidad. Editorial Nueva imagen S.A. de C.V. México. 1999. pags. 37, 38, 39. 59, 60, 61,90, 91, 92, 107, 108, 127, 170, 171, 188, 189, 195, 207, 208, 209.

FLORES, Meyer Marco; Temas selectos de Matemáticas. Editorial Progreso S.A. México. 1981. pags. 236, 237, 238, 246, 247, 249, 250, 251.

SALAZAR, Poot Lucio; Estadística. NAFINSA. México. Pags.1, 6, 7, 9,11, 12.

MARMOLEJO, G. Martín; Inversiones, IMEF. México. 1987.