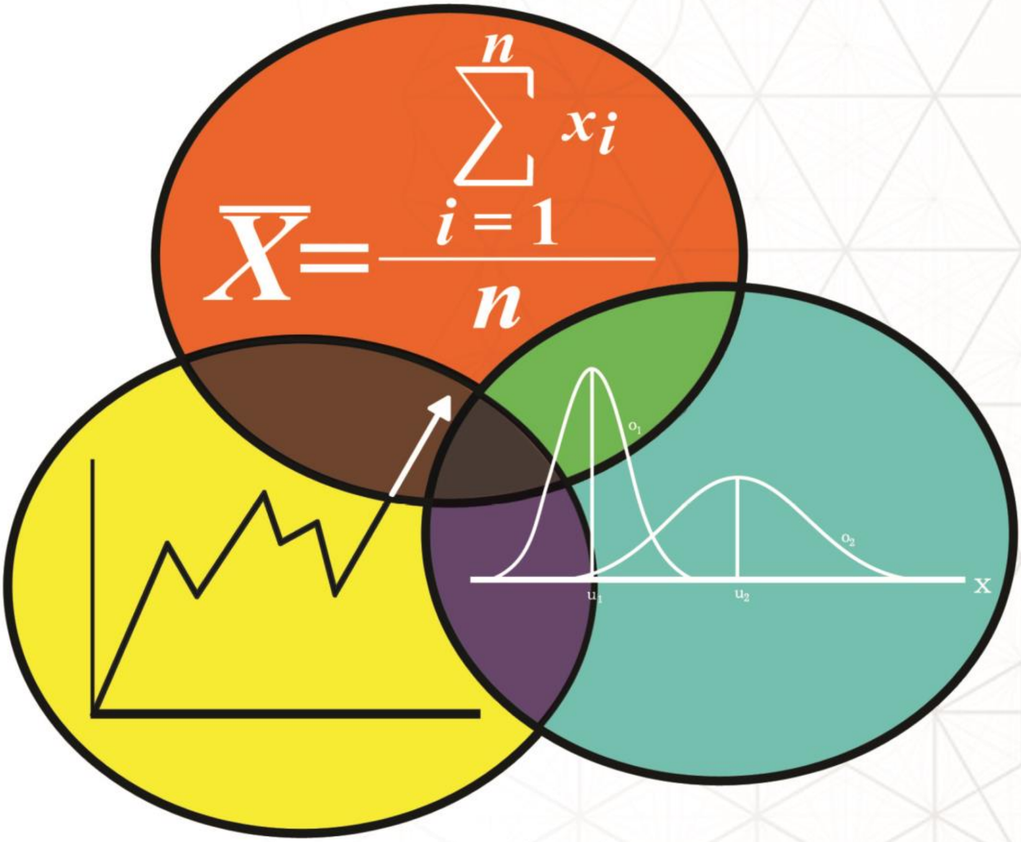
ESTADÍSTICA

APLICADA A LA

ADMINISTRACIÓN Y LA ECONOMÍA

Primera Edición



Mauro C. Tapia T.
Elaine R. Jijón G.

Estadística aplicada a la administración y la economía

Estadística Aplicada a la Administración y la Economía

PRIMERA EDICIÓN

MAURO C. TAPIA TORAL
ELAINE R. JIJÓN GORDILLO
Catedráticos de la Facultad de Ciencias Económicas
Universidad de Guayaquil
Ecuador

REVISORES TÉCNICOS

IRAMARÚ HERRERA
Docente e Investigadora
Universidad de Carabobo
Venezuela

JOHANDRY LÓPEZ
Licenciada en Ciencias Estadísticas
Universidad Central de Venezuela (UCV)
Universidad Católica Andrés Bello (UCAB)
Venezuela

Estadística aplicada a la administración y la economía

Centro de Investigación y Desarrollo Ecuador. Cdla. Martina Mz. 1 V. 4 - Guayaquil, Ecuador. Tel.: 00593 4 2037524 http. :/www.cidecuador.com
ISBN: 978-9942-759-52-8
Impreso y hecho en Ecuador
Diseño y Tipografía: Lic. Pedro Naranjo Bajaña
Ing. Duberli Jiménez Cabrera
Fecha de Publicación: Enero 2018

CIDE EDITORIAL
Cod. 9942-8632

Guayaquil - Ecuador

AGRADECIMIENTO

A Dios todopoderoso quien por medio de Él, todo es posible.

A mi madre quien con su fiel dedicación generó las huellas del camino.

A mi difunto padre por su aporte en lo académico y profesional.

Mauro C. Tapia T.

AGRADECIMIENTO

Al Ser Supremo por la bendición de disfrutar la culminación de un propósito más en mi vida.

A mi madre por su persistencia perpetua y bondad infinita que no se agota ni en los momentos más duros.

A mi padre por inculcarme valores de honradez e integridad.

A mis alumnos, por su constante inquietud y creatividad, fuente de inspiración y mejora contínua para el docente.

A las autoridades de la Facultad de Ciencias Económicas de la Universidad de Guayaquil por el fomento a la cultura de la escritura científica, factor importante en los procesos de construcción del conocimiento, a través de la expresión escrita.

Elaine R. Jijón G.

DEDICATORIA

A mi querido hijo, pilar fundamental de fuerza, motivación y energía para seguir generando nuevos propósitos en la vida.
A mi querida madre, fuente de apoyo e ímpetu y sagacidad.

Mauro C. Tapia T.

DEDICATORIA

A mi hija Lolita, la primera y sublime creación de mi vida.
Dedico mi segunda obra como testimonio de que la constancia vence las barreras y dificultades en el camino a los sueños y puede modificar el sino desventurado de un acontecimiento.

Elaine R. Jijón G.

ACERCA DE LOS AUTORES

Mauro Carlos Tapia Toral

Magister en Finanzas y Proyectos Corporativos de la Universidad de Guayaquil. Diploma Superior en Docencia Universitaria de Universidad Técnica de Machala. Economista con mención en Gestión Empresarial. Especialización Finanzas de la Escuela Superior Politécnica del Litoral. Contador Público Autorizado de la Universidad Estatal de Milagro. Docente de la Universidad de Guayaquil.

Elaine Raquel Jijón Gordillo

Master of Business Administration-MBA at UQAM Université du Québec à Montréal. Magister en Administración de Empresas, ESPOL. Especialista en Gestión Financiera en Instituciones Públicas y Privadas. Docente en el área de Métodos Cuantitativos y Administración y Control de Calidad. Economista con especialización en Administración de Empresas y Gestión de Procesos.

Contenido sintetizado

Unidad 1
La estadística y sus aplicaciones

Unidad 2
Agrupación y presentación de datos estadísticos

Unidad 3
Estadística sumaria, principales medidas y sus aplicaciones

Unidad 4
Números índices

Unidad 5
Teoría de conjuntos, coordinatoria y de probabilidades

Unidad 6
Variables aleatorias y distribución conjunta

Índice general

Unidad 1
La estadística y sus aplicaciones

Unidad 2
Agrupación y presentación de datos estadísticos

Unidad 3

Estadística sumaria, principales medidas y sus aplicaciones

Unidad 4
Números índices

Unidad 5
Teoría de conjuntos, coordinatoria y de probabilidades

Unidad 6
Variables aleatorias y distribución conjunta

Apéndice 7

Índice de tablas

Unidad 2
Agrupación y presentación de datos estadísticos

Unidad 3
Estadística sumaria, principales medidas y sus aplicaciones

Unidad 4
Números índices

Unidad 5
Teoría de conjuntos, coordinatoria y de probabilidades

Unidad 6
Variables aleatorias y distribución conjunta

PRÓLOGO

Los creadores del presente trabajo, utilizando un criterio sencillo, dinámico y espontáneo, han desarrollado un esquema ordenado para explicar cómo la Estadística siendo un subconjunto de las Ciencias Matemáticas, permite utilizar herramientas de análisis de datos para la toma de decisiones. Dada su estructura, el texto ha sido diseñado para ser utilizado a nivel de pequeños, medianos y grandes empresarios, así como también estudiantes de licenciatura de la carrera de Economía de las diferentes universidades del país.

Los autores introducen retrospectivamente, en cada capítulo, los conceptos teóricos de la estadística, para posteriormente de forma didáctica, continuar con ejercicios prácticos de presentación de los datos, análisis de las medidas de tendencia central, números índices, teoría de conjuntos y probabilidades, y así involucrarnos con las variables aleatorias su distribución conjunta, desarrollando en cada sección, ejemplos derivados del conocimiento y la experiencia académica, acumulados en el desarrollo de la cátedra universitaria.

Iniciativas como éstas, que toman en cuenta la habilidad del docente de generar la aplicación práctica de fundamentos teóricos, merecen nuestro apoyo académico para que continúen creando nuevas propuestas metodológicas, en esta área cuantitativa, tan útil para el desarrollo de todas las ramas del conocimiento científico.

Econ. César Saltos Veliz, Msc.
Candidato a Phd. en Economía de la Universidad La Molina de Perú
Jefe de Cátedra área Econometría
Facultad de Ciencias Económicas
Universidad de Guayaquil

Unidad 1
LA ESTADÍSTICA Y SUS APLICACIONES

-KARL PEARSON -

Historiador, escribió sobre folklore, fue un socialista convencido, abogado, matemático aplicado, biómetra, estadístico, maestro y biógrafo. Pero sin duda su contribución más importante es al nacimiento de la Estadística Aplicada.

”Hasta que los fenómenos de cualquier rama del conocimiento no hayan sido sometidos a medida y número, no se puede decir que se trate de una ciencia”.

CONTENIDO DE LA UNIDAD

INTRODUCCIÓN

En este apartado se mostrará en síntesis la definición de la Estadística como ciencia exacta y su desarrollo a través de la historia, también se explicaran los principales conceptos relacionados a la nomenclatura estadística. Además se presentaran los programas más utilizados para el desarrollo de cálculos y organización de datos muestrales.

OBJETIVOS DE APRENDIZAJE

Al concluir el capítulo, será capaz de:

- Conocer el origen y evolución de las ciencias Estadísticas en el Ecuador
- Definir conceptos básicos de las ciencias Estadísticas.
- Diferenciar los programas Estadísticos.
- Describir las variables estadísticas
- Diferenciar la Estadística Descriptiva de la Estadítica utilizada en el análisis de datos.

Hay que resaltar la importancia que ha tenido la utilización de los métodos estadísticos al momento de realizar ensayos para la búsqueda de resultados. En nuestro país, un sistema de institucionales gubernamentales, dedicadas a desarrollar y controlar la información estadística del país, entre ellas el Instituto Nacional de Estadísticas y Censos (INEC), donde se realizan estudios demográficos, políticos, sociales y económicos.

1.1 Definición de la estadística y su utilidad

1.1.1 Definición de estadística

Todas las ramas del conocimiento humano, en el proceso de investigación, utilizan a la ciencia estadística para realizar experimentos, partiendo de la obtención de datos, principalmente numéricos los cuales utilizan inicialmente, una metodología descriptiva.

La primera fase del proceso estadístico es la recolección de datos, es decir producir información que permita ser analizada. Este primer camino nos lleva a contar con la materia prima de la investigación, sea a través de encuestas o entrevistas, la misma que luego será ordenada u organizada, descartando aquellos datos que no cumplen con los parámetros de calidad o el propósito del trabajo.

El proceso de tabulación permite obtener información depurada que sigue a la fase de presentación de los datos, donde éstos se muestran en tablas y gráficos. Con lo cual describimos el conjunto de datos observados a través del ordenamiento de dicha información.

Si requerimos una caracterización más sucinta del objeto de investigación, elaboramos índices que son el resumen de los datos en un solo indicador y utilizamos las medidas de tendencia central o promedios, las medidas de dispersión, las medidas de asimetría o las medidas no centrales, con lo cual podemos explicar un conjunto de datos, sólo hablando de este indicador.

Cuando se trata de análisis más complejos en los cuales no podemos tener acceso a todos los datos, utilizamos una muestra seleccionada de esa población, para a través del análisis de sus resultados, inferir conclusiones respecto a toda la población. En este último caso estamos utilizando un método inferencial, para el análisis.

Haciendo un análisis más detallado, observamos que el concepto de estadística incluye los siguientes elementos:

1.-Es una ciencia (involucra métodos científicos).

Estudia la recolección, organización, presentación, resumen, análisis e interpretación (fases del proceso estadístico):

- De datos.
- Que pueden ser traducidos a números.
- Con el fin de tomar decisiones.

2.- Permite, a través del análisis de una muestra, sacar conclusiones respecto al comportamiento de una población.

Entre las definiciones más aceptadas de estadística, tenemos:

> "La estadística es la ciencia que estudia los métodos científicos para acumular, establecer, sintetizar y analizar datos, así como para lograr conclusiones válidas y tomar medidas razonables fundadas en tal análisis, esta ciencia con base matemática investiga condiciones regulares en fenómenos de tipo aleatorio. Es transversal a una ancha complejidad de disciplinas, partiendo con la física y abarcando las ciencias sociales, ciencias de la salud, de control de calidad y es también usada para la toma de decisiones en áreas comerciales, negocios y organismos gubernamentales."

Al separar este concepto en dos partes, se puede concluir que la estadística se divide en descriptiva e inferencial.

Como su nombre lo dice, la estadística descriptiva tiene como propósito describir un fenómeno o un conjunto de datos, para ello se vale de tablas, gráficos (presentación), o indicadores (resumen), como promedios, dispersión y percentiles.

La Estadística inferencial en cambio, usando métodos probabilísticos, podrá a través de la selección de una muestra, sacar conclusiones de toda la población.

Según la etimología de la palabra, estadística proviene del vocablo alemán Statistik que se trascribe de manera entendible como "La ciencia del Estado" (David Ruíz Muñoz, Universidad Pablo de Olavide, 2008). En el siglo XIX, el británico Sir

John Sinclair (1754-1835) expone un sencillo término de la ciencia estadística el cual significa recolectar y clasificar datos.

Transcurriendo el siglo XX, el desarrollo de instrumentos precisos para cuestiones de salud pública (epidemiología, bioestadística, etc.) y planes económicos-sociales (tasa de desempleo, econometría, etc.) precisó de avances esenciales en las prácticas estadísticas.

Hoy en día, el uso de la estadística se ha desarrollado más allá de sus orígenes como una asistencia al Estado o al gobierno. Personas y organismos usan la estadística para hallar información y tomar decisiones en ciencias médicas, naturales, sociales, negocios y otras áreas. Esta ciencia empírica es razonada generalmente no como un simple modelo de las matemáticas sino como una ciencia diferente.

1.1.2 Utilización básica de las ciencias estadísticas

En el ámbito de acción de la estadística descriptiva, la información cuantitiva y cualitativa obtenida a través del proceso de recolección de datos, permite al administrador presentar en forma resumida las características de un fenómeno económico o social, de allí que su nombre de "estadística descriptiva" se deriva de la capacidad de explicar o definir el comportamiento de una variable, a través de un conjunto de tablas y gráficos o resumir la evolución de un fenómeno social con base en un índice o indicador, sea un promedio, una dispersión, un coeficiente de variación, en ventas, el consumo, producto interno bruto, número de enfermedades, tasa de mortalidad, tendencia de crecimiento de la pobreza, entre otros.

La estadística inferencial, en cambio a través del análisis de una muestra, permite sacar conclusiones de la población, es decir del todo, como por ejemplo, en el caso de un exit poll, estimar quién es el candidato ganador de una elección, solo analizando un determinado número de electores. Determinar a través de una prueba en un laboratorio si se debe utilizar un medicamento para curar una enfermedad considerando un margen de confianza y un riesgo medible, darle al administrador la confianza de que una hipótesis puede ser rechazada o aceptada para la toma de una decisión.

La estadística inferencial formó un número formidable de “materiales” de los métodos estadísticos que usan los expertos de la estadística. Los métodos estadísticos se plantean para contribuir al asunto de ejecutar juicios científicos ante la incertidumbre y la variación.

Hoy en día, la fabricación de los productos farmacéuticos, el software para computadoras, el desarrollo de productos alimenticios, las fuentes de energía y muchas otras áreas, implica la recolección de información o datos científicos. Por supuesto que la preparación de datos no es algo desconocido y no es lo único importante. Los datos, a lo largo del tiempo, se han conseguido, abreviado, alcanzado y acumulado para su análisis cuidadoso. No obstante, hay un contraste profundo entre la recolección de información científica y la estadística inferencial. Esta última ha tomado cuidado respectivo en décadas recientes.

Una de las situaciones que impulsó el uso de la estadística dentro del proceso de industrialización, es que la densidad de producto de un compuesto específico no siempre será la misma. Si un paso es discontinuo en vez de continuo, la densidad de compuesto no sólo se renovará entre los lotes que salen de la línea de producción (cambio de un lote a otro), sino igualmente dentro de los propios lotes. Los métodos estadísticos se usan para estudiar datos de procesos como el primero. El objetivo de esto, es obtener una excelente orientación en relación de cuáles cambios corresponde realizar en el proceso para optimizar su calidad.

Es así como, en cualquier campo del conocimiento científico y cualquier proceso industrial o de fabricación, se utiliza la estadística para la toma de decisiones a través del análisis de la información obtenida.

1.2 Nociones básicas en estadística

A continuación se muestran algunos términos utilizados en la estadística (Instituto Nacional de Estadísticas e Informática, 2006):

Estadística descriptiva: Método para establecer, abreviar y mostrar datos de manera informativa.

Estadística inferencial: Métodos utilizados para establecer una cualidad de una población con base en la búsqueda de una muestra.

Población: Conjunto de elementos o características a los que se les estudia mediante un ensayo.

Individuo: Cada uno de los elementos que conforman el grupo de la población.

Muestra: Subconjunto específico de la población.

Modalidad: Cada una de las eventos o etapas diferentes de una variable estadística.

Experimento: Es un método de exploración mediante el cual se establece la incidencia de variables independientes sobre la variable dependiente.

Tipos de experimentos:

Experimento determinista: El resultado se encuentra establecido por algún teorema.

Experimento aleatorio: No se puede producir un resultado esperado ya que está sujeto al azar.

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

Tipos de variables:

Variable estadística: Particularidad propia del individuo u objeto del estudio estadístico.

Variables cualitativas: Las características no son cuantificables, se basan en cualidades como: el tipo de raza, el estado civil, entre otros.

Variable cualitativa dicotómica: Cuando sólo alcanzan escoger dos opciones posibles como sí y no, hombre y mujer o son politómicas cuando alcanzan adquirir tres o más opciones de valores a escoger.

Variable cualitativa ordinal: La variable consigue brindar distintos valores ordenados alcanzando una escala establecida, aunque no es obligatorio que el intervalo entre exactitudes sea uniforme, por ejemplo: grave, moderado, leve.

Variables cualitativas nominal: En esta variable los valores no consiguen ser sometidos a una razón de orden como por ejemplo: el lugar de residencia o los colores.

Variables cuantitativas: Son de características cuantificables, medibles o numéricas, como número de lotería, horas del día, entre otros.

Variable cuantitativa discreta: Esta variable presenta alejamientos o interrupciones en la escala de valores que puede escoger. Estas dispersiones o interrupciones indican la ausencia de valores entre los diferentes valores determinados que la variable pueda adjudicarse. Un ejemplo: es el número de hijos en un hogar.

Variable cuantitativa contínua: Esta variable puede conseguir cualquier valor entre de un intervalo especificado de valores. Por ejemplo la altura o el peso, que simplemente es restringido por la precisión del aparato medidor, en el supuesto permiten que constantemente exista un valor entre dos cualesquiera.

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

1.2.1 Ejemplos de definiciones básicas

Población: Los estudiantes de la Universidad de Guayaquil.

Muestra: Los estudiantes de la Facultad de Economía de la Universidad de Guayaquil.

Variable cualitativa:

- Sexo de los estudiantes.
- Equipo de fútbol preferido por los estudiantes.
- Estado civil de los estudiantes.

Variable cuantitativa:

- Edad de los estudiantes de la Facultad de Economía.
- Salario de los contadores públicos al inicio de su carrera.
- Nota promedio de un egresado de una universidad pública.

Variable discreta:

- Número de accidentes de tránsito en la provincia del Guayas, ocurridos en el año como consecuencia de exceder el límite de velocidad.
- Número de electores en la provincia de Manabí.

Variable contínua:

- Peso de un niño al nacer.
- Área promedio de un lote de terreno en una urbanización.
- Ingreso mensual promedio de los gerentes de una institución financiera.

Variable cuantitativa: Edad promedio de los habitantes en Vilcabamba.

Variable cualitativa: Plato típico preferido en Vilcabamba.

Variable discreta: Número de habitantes en Vilcabamba.

Variable contínua: Ingresos mensuales de los hostales en Vilcabamba.

Variable cualitativa ordinal: Calificación del servicio de atención al turista en la ciudad:

- Deficiente, regular, bueno, muy bueno, excelente.

1.2.2 Población y muestra

A fin de evitar resultados errados al sacar conclusiones para la población, basándose en una muestra, la muestra utilizada en una investigación científica, debe cumplir con las siguientes condiciones:

1. **La muestra debe ser aleatoria**

Esto quiere decir que los elementos elegidos en la muestra deben ser seleccionados al azar (mediante un sorteo) o empleando cualquier método que asegure que no existe sesgo en su elección. Las personas u objetos participantes de la población deben tener la misma probabilidad de ser elegidos.

2. **La muestra debe tener el tamaño mínimo adecuado**

Esto indica que el número de datos elegidos para la muestra debe ser del tamaño que permita minimizar el error de estimación. El cálculo del tamaño de muestra requerido dependiendo si la población es finita o infinita será analizado posteriormente y es producto de una fórmula estadística que incluye: margen de error aceptable, desviación estándar de la población, nivel de confianza y en caso de tenerlo, tamaño de la población

3. **La muestra debe ser representativa de la población**

Es decir la muestra debe reproducir las características de la población en su misma estructura o proporción, por ejemplo si en la población existe 30% de personas de la costa y 70% de personas de la Sierra, debe mantenerse esa relación en la distribución de la muestra.

1.2.3 Parámetros y estimadores

Se llama parámetro a un indicador relacionado con una población, por ejemplo las exportaciones promedio del Ecuador es un parámetro.

Se denomina estadígrafo a un indicador que representa los datos de una muestra, de allí que las exportaciones promedio de una empresa bananera, es un estadígrafo.

Dado que inferencia estadística es la conclusión que hacemos acerca de la población basado en los datos de una muestra, los indicadores de la muestra, se convierten en “estimadores”, de los parámetros de la población.

Un estimador es cualquier índice calculado con los datos de la muestra, con el cual podemos llegar a aproximar el verdadero valor poblacional. Ejemplo: en una muestra de consumidores consultamos su preferencia respecto al consumo del sabor de un helado, con estos datos, calculamos la proporción muestral de preferencia al sabor de un helado, este indicador p, es un estimador de la verdadera preferencia poblacional acerca del sabor.

1.2.4 Tipos de muestreo

El muestreo es la selección de una muestra y su tratamiento para el análisis de los resultados de la población.

Si la muestra no es representativa de la población los resultados pueden ser errados para hacer una conclusión respecto del todo.

Los tipos de muestreo son:

Muestreo aleatorio simple. Todos los elementos que conforman la población tiene la misma oportunidad de ser seleccionados en la muestra. Se utilizan números aleatorios para selección o el método del sorteo o elección dentro de una urna de todos los elementos.

Muestreo aleatorio estratificado. Se elige la muestra segmentando la población en estratos, se calcula su porcentaje y esta misma estructura se aplica a la muestra. Se aplica para cuando la población está dividida en segmentos. En los estratos los individuos al interior son homogéneos, pero entre cada estrato si hay heterogeneidad.

Muestreo sistemático. Este método se utiliza cuando la población esta ordenada físicamente pero no se haya un registro ordenado para realizar un muestreo aleatorio simple.

Muestreo por conglomerado. Se utiliza cuando los elementos de un grupo son heterogéneos al interior, pero como grupo son homogéneos. Es decir, dentro de los elementos de cada grupo hay diversidad, pero entre cada grupo hay similitud.

Se diferencia del muestreo estratificado, porque el investigador puede seleccionar los conglomerados con los que trabajará.

1.3 Estudios estadísticos del Ecuador

1.3.1 Antecedentes

La importancia de la estadística como ciencia en el Ecuador se aprecia través de la historia, si con frecuencia se indica el uso distante de las numeraciones aborígenes, mediante los quipus, así como los padrones romanos, en el presente territorio ecuatoriano dicha disciplina adquiere fuerza con la colonización hispánica, que solicita el uso de la agrupación cuantitativa como una necesidad primordial. Constituir una organización colonial efectiva, posterior a la conquista y cambio del mundo aborigen. Desde los principios de la Colonia el tributo se convirtió en una parte fundamental del desarrollo político de administración del territorio ecuatoriano (Instituto Nacional de Estadísticas y Censos (INEC), 2015).

1.3.2 Antecedentes históricos y la proto-estadística

Los empadronamientos o censos en la antigua colectividad colonial fueron la base primordial para la imposición de impuestos, la clasificación del status étnico y consecutivamente, para la representación política y el servicio militar. Fueron asimismo por propio derecho los mayores acontecimientos, como lo manifestaron las rebeliones indígenas que acompañaron a los censos de aproximadamente en los años 1764, 1765 y 1780.

La demografía ecuatoriana republicana: 1825-1950

Con la incorporación de la flamante República de la Gran Colombiana Bolivariana, se comienza la actividad de la enumeración censal, nombrando a los documentos resultantes "Estados generales de población y noticias estadísticas, 1825-1827", como los distritos de Cuenca y Loja, levantados de consentimiento con lo dispuesto por el artículo 8° de la "Ley sobre la organización y régimen político y económico de los departamentos y provincias", (Instituto Nacional de Estadísticas y Censos (INEC), 2015) del 11 de marzo de 1825 y de acuerdo con el cuestionario elaborado en el decreto ejecutivo del 4 de octubre de 1825.

Con fundación de la República, cambia la expresión administrativo institucional con dependencia al colonial hispánico y se utiliza llanamente el nombre **"censo de la población"** a modo de **"noticias estadísticas"** a nivel de cuadras, parroquias, cantones, provincias, departamentos. Con la institución de la República del Ecuador, sus iniciales presidentes, Juan José Flores y Vicente Rocafuerte, prolongan con el interés de establecer una enumeración mejorada y al día de toda la población ecuatoriana. Estos personajes personalmente, obligaron entre 1838 y 1840 a los intendentes de las jurisdicciones y gobernadores de las provincias a realizar nuevos empadronamientos.

Sin embargo, los promotores del censo fueron en su mayoría funcionarios capaces y minuciosos. La ejecución del Primer Censo Nacional de Población (1950) simboliza el inicio de una nueva fase estadística en el Ecuador. La integración mundial determina la asimilación de la metodología divulgada por Naciones Unidas y el gobierno interamericano como factores de homogenización técnica mundial. Si de este modo, se prepara la etapa propiamente estadística porque el país ya puede detallar con series contínuas de indicadores.

Ilustración 1.1 Encabezado Diario El Telégrafo
Fuente: Archivo El Telégrafo; publicada el día miércoles
29 de Noviembre de 1950.

De tal forma, con el lapso del tiempo se comienza en la elaboración de multitudinarios estudios demográficos, de salud, económicos, transporte, educativos e infraestructura los cuales se perfeccionan con la enseñanza de la estadística en las universidades nacionales.

En el Ecuador, en la década de 1940, pese a la divulgación de diversas estadísticas, adquiere relevancia en la opinión pública la idea de la presencia de un vacío de investigación estadística, ya que la propagación de múltiples causantes de datos estadísticos, la desaparición de capacitación, la necesidad de detallar con más indicadores estadísticos y el dominio y presiones de los entidades internacionales, establecieron la necesidad de contar con una entidad competitiva a fin de generar la estadística ecuatoriana.

Primer censo del Ecuador

La formación del Primer Censo de Población personifica una tarea compleja para el Estado ecuatoriano, porque solicitaba la conformación de una estructura institucional que dispusiera de una logística y el financiamiento en el argumento de un proceso de homogenización corporativo internacional.

Se asistía de realizar un padrón con un contenido diferente, mucho más amplio, de variables que debían compilarse en el campo. En este argumento, se trataba de un asunto inédito porque era metodológicamente nuevo pero no se podía solicitar, como lo hicieron sus regentes, con una visión de la década de los años cuarenta, la ausencia de prácticas estadísticas anteriores cuando en esos mismo años se efectuó el empadronamiento de las ciudades de Quito y Guayaquil, y se detallaba con las cifras contínuas de los centros del Registro Civil con un nivel demográfico.

Último censo del año 2010

El día domingo 28 de noviembre de 2010, con respaldo de 500 mil personas, entre estudiantes, profesores, parte de las Fuerzas Armadas, recursos de la Policía Nacional y miembros del INEC se extendieron por todas las zonas demográficas del territorio nacional con una misión común: hacer del VII Censo de Población y VI de Vivienda, un movimiento civil e incluyente exitoso. El informe del censo recogió 71 preguntas concentradas en el domicilio, el hogar, la migración, ingresos familiares y la población. La indagación fue levantada por 361 mil alumnos de bachillerato y profesores. La acción censal se ejecutó entre las 07:00 y las 17:00, período en el que las ciudades lucieron vacías pues los habitantes, se perseveraron en sus hogares a la atención de ser censados, los medios de transporte urbano e interprovincial fueron limitados y las fronteras, mercados y otras entidades siguieron cerrados.

Así, en la noche del 27 de noviembre, los principales censados fueron las personas sin vivienda ya que se brindó especial importancia al levantamiento de búsqueda de las personas con discapacidad o en situación de calle. El argumento de la migración también fue apreciable, por única vez en la historia censal se efectuó un proceso a partir de los pueblos y nacionalidades para levantar participativamente las preguntas de auto identificación, mediante las cuales se cuestionaba a qué grupo étnico, pueblo o nacionalidad pertenecía, apoyando así el derecho de las personas a disponer de forma libre y voluntaria su participación a una etnia, nacionalidad o pueblo indígena, afrodescendiente o montubio.

El INEC alcanzó con el censo a los territorios más excluidos del país. Los empadronadores transitaron varios kilómetros para llegar incluso al último rincón de la zona geografía nacional en una labor que para la zona rural tomó siete días, fue desde el 29 de noviembre al 5 de diciembre de 2010 (Instituto Nacional de Estadísticas y Censos (INEC), 2015).

1.4 Programas informáticos para el desarrollo y análisis de datos estadístico

Entre de los programas estadísticos más manipulados están: SAS, SPSS, STATISTICA, STATA y, últimamente está, MATLAB. A continuación se mostrará una breve introducción algunos programas de mayor uso estadístico.

SPSS (Statistical Package for The Social Sciences)

Ilustración 1.2 Programa Informático SPSS
Fuente: www.ibm.com

Este programa, es uno de los más divulgados, también perfeccionado por la Universidad Norte Americana de Chicago. Es un programa estadístico, de automatismo ordinario, que forma procedimientos descriptivos y gráficos asimétricos de muy alta resolución, de tal modo que se aprovecha en el apoyo al análisis de datos. El programa se vale para desarrollar temas como: Pronósticos con series de tiempo, métodos de investigación, finanzas, métodos cuantitativos, inferencia estadística, segmentación de mercados, análisis multivariado, métodos multivariados y otros más. Otra de las características es que se pueden realizar estudios exploratorios con un punto de vista gráfico, de la misma forma se manipulan para ejecutar análisis estadístico simple y/o avanzado (IBM Corp., 2017).

SAS (Statistical Analysis System)

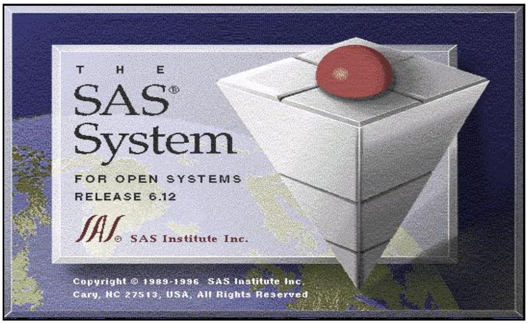


Ilustración 1.3 Programa Informático SAS
Fuente: www.sas.com

Consigue amplias posibilidades de operaciones estadísticos (regresión variada con posibilidades diagnósticas, métodos multivariados, análisis con riesgos proporcionales y regresión lineal) y permite sistematizaciones exactas para tablas Z, T, Gamma y contiene poderosas resoluciones gráficas.

- Los procedimientos pueden aprovecharse de una sola realización:
- Los resultados se pueden guardarse como simples archivos para luego utilizarse como entradas para futuras prácticas.
- Es exclusivamente útil en la gestión y organización de datos, también en la redacción de informes.
- Los procedimientos pueden tener algunas opciones por lo cual, se debe examinar meticulosamente el manual antes de escoger la opción deseada.

SAS brinda la flexibilidad para individualizar el manejo y estudio de datos, sin embargo el principal problema, es que no es fácil aprender a usarlo (NC State University, 2017).

STATISTICA

Ilustración 1.4 Programa Informático STATISTICA
Fuente: www.statsoft.com

Statistica (no obstante la marca está reconocida como STATISTICA) es un sistema de programación estadístico utilizado en investigación, análisis de datos y en el contorno empresarial. Lo creó STATSOFT, compañía que lo mejora y conserva. STATSOFT se perfeccionó en 1984 de un acuerdo entre un conjunto de profesores universitarios y científicos (Statsoft, 2017).

Finalmente, en 1991, se arrojó al mercado de programación computarizada la primera versión de STATISTICA. El programa contiene varios módulos. El primordial de ellos es la Base, que efectúa análisis con las técnicas estadísticas más comunes. Éste puede completarse con otros módulos específicos tales como:

- Advanced: que contiene las técnicas multivalentes y modelos progresados de regresión lineal y no lineal.

- QC: posee las técnicas de control de calidad y los estudios de procesos (distribuciones normales, Gage R&R, Weibull) y de diseño experimental.

MATLAB

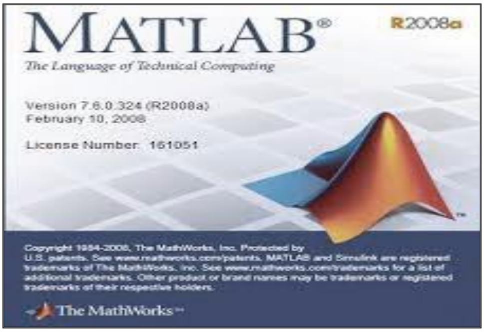


Ilustración 1.5 Programa Informático MATLAB
Fuente: www.mathworks.com

Se origina como un recurso para la necesidad de las mejores y únicas herramientas de cálculo para solucionar problemas de cálculo complicados, en los que es necesario beneficiarse de las amplias capacidades de proceso de datos.

Éste es un desarrollo de aplicación totalmente integrada, dirigida para realizar proyectos en donde se encuentren involucrados grandes cálculos matemáticos y la observación gráfica de los mismos.

MATLAB compone análisis numérico, desarrollo de cálculo matricial proceso de funciones y visualización gráfica en un ambiente completo en el cual los problemas y sus resultados son expresados de la misma manera en que se representarían tradicionalmente, sin hacer uso de la programación tradicional.

Para culminar se puede decir que MATLAB es un programa de alto nivel y un entorno participativo para el cálculo, visualización y programación. El lenguaje de la programación, las herramientas y funciones algebraicas integradas que acceden explorar varios enfoques y llegar a una solución más rápida (MathWorks , 2017).

Comparación entre los dos paquetes de programación estadística SAS y SPSS:

Facilidad: SPSS es mucho más sencillo que el de SAS. No obstante, SAS, una vez aprendido el manejo de sus herramientas de programación, es más agilizado que SPSS.

Formación: SPSS al ser más fácil, no sujeta al usuario a extensos procesos formativos, sino que, en un plazo de tiempo más corto que SAS, puede realizar complicadas operaciones de análisis sin esfuerzo.

Precio: El precio de venta el mercado de programas computacionales, es que SPSS tiene un menor precio que su contraparte el programa SAS, por lo tanto su demanda se enfoca a profesionales o desarrolladores de estudios particulares.

Robustez: El programa estadístico SAS, al poseer su propio sistema operativo, alcanza la iniciativa del sistema, no permitiendo que una acción paralice el trabajo del operario, lo contrario de SPSS.

Rendimiento: Para el programa SPSS, su sistema operativo permite una mejor solución, pero para las compañías que puedan aceptar un desembolso de efectivo sin necesidad de retorno a corto plazo, SAS es más rentable.

1.5 Ejercicios propuestos

1.- Elabore una lista de las ciencias que utilizan a la estadística para el desarrollo de experimentos. Proporcione ejemplos de cada uno.

2.- Ubique las siguientes variables en las tablas de clasificación:

- Salario.
- Número de personas en el salón.
- Género.
- Volumen de ventas de reproductores MP3.
- Preferencia por los refrescos.
- Descripción de personas a su alrededor.
- Temperatura.
- Resultados de la Lotería Nacional.
- Lugar que ocupa un estudiante en clase.
- Calificaciones de un profesor de estadística.
- Cantidad de computadoras domésticas.
- Actividades en común con sus compañeros.
- Horas de estudio de estadística a la semana.

- Ventas del departamento de ropa.
- Preferencia de tipo de mascotas.
- Lista de útiles escolares.
- Consumo de comidas rápidas.
- Tipos de dispositivos móviles.

	VARIABLE
CUALITATIVA	
CUANTITATIVA	

3.- ¿Cuál es el nivel de medida de cada variable? (cualitativa, cuantitativa)

a) Coeficiente intelectual de los estudiantes.
b) La distancia que viajan las ballenas.
c) Las calificaciones en el primer examen de estadística.
d) Clasificación de personas por fecha de nacimiento en el salón de clase.
e) Una agrupación de estudiantes que cursan último grado.
f) Tiempo que los alumnos estudian a la semana.

4.- Ingrese a la página de Internet del INEC. Halle las siguientes ponderaciones y luego indique el tipo de variable de que se trata.

- Número de personas en a nivel nacional.
- Número de personas que habitan en Guayas, Pichincha y Cuenca.
- Porcentajes de las principales etnias o grupos étnicos en el país.
- Género de los habitantes.
- Promedio de personas por hogar.
- Promedio de ingreso per cápita.

5.- Elabore un cuadro, detallando qué tipos de variables afectan los siguientes temas de interés económico.

- Alza de pasajes del transporte urbano.
- Incremento de la canasta básica.
- Alza del precio del barril de crudo ecuatoriano.
- Mejoras del sector productivo en el Ecuador.

- Eliminación de subsidio al gas licuado de petróleo.
- Variación en la tasa de desempleo.

6.- Hallar los datos estadísticos del censo realizado en el año 2010. Determinar los resultados de las provincias de Guayas y Pichincha.

7.- Investigue la información referente a los censos realizados décadas anteriores y luego con dicha información y el apoyo de el programa estadístico de su preferencia, elabore un gráfico del incremento poblacional del país.

8.- Investigar como mínimo cuatro ejemplos de fenómenos económicos, considerados dentro del campo de la investigación estadística.

9.- Investigue qué otras instituciones públicas o privadas a nivel nacional realizan estudios estadísticos o relacionados a la manipulación de datos estadísticos.

10.- En la página de internet del INEC. Hallar las proyecciones de las siguientes variables:

Población.

Género.

11.-Investigue: ¿Por qué es útil la estadística en la actividad para la cual se está preparando? Prepare argumentos y expóngalo en clase.

12.- Investigar, el uso de la Estadística en las siguientes ramas de la ciencia y proporcione ejemplos:

- Medicina.
- Psicología.
- Industrial.
- Social.
- Educacional.

Unidad 2

AGRUPACIÓN Y PRESENTACIÓN DE DATOS ESTADÍSTICOS

“La lectura e interpretación de los gráficos es una ciencia que busca inútilmente lo que el saber consigue.”
-ANDRÉ KOSTOLANY-
Especulador y experto en bolsa reconocido mundialmente. Trabajó la mayor parte de su vida en Francia y Alemania.

CONTENIDO DE LA UNIDAD

INTRODUCCIÓN

En este capítulo se realiza la introducción de la presentación de la información obtenidas mediante un muestreo, tanto textual (en forma de texto); Cuadros y Gráficas. La elaboración de informes que permite presentar los datos obtenidos de la investigación en tablas y gráficas, a través de la cual se muestra el comportamiento de la variable

OBJETIVOS DE APRENDIZAJE

Al culminar el capítulo, será capaz de:

- Identificar los métodos de recolección de datos y las fuentes de obtener información.
- Estructurar tablas de distribución de frecuencias para variables cuantitativas y su análisis respectivo.
- Ordenar y agrupar los datos cualitativos para la elaboración de una tabla de distribución de frecuencias.
- Representar un cuadro de frecuencias como una gráfica de barras o de pastel.
- Interpretación de datos cuantitativos en una distribución de frecuencias.

La importancia de los métodos de recolección de datos radica en que permite organizar el levantamiento de información requerido para mostrar, la totalidad de los elementos que componen la muestra, población o universo bajo estudio. Cuando se realiza un censo, se debe cumplir las condiciones de universalidad (empadronar a todos los elementos una población) y simultaneidad (efectuarse en un momento determinado). El informe de la recopilación de datos muestra en resumen el comportamiento de todas las unidades muéstrales, como es el caso de los censos de población, económicos, agropecuarios, etc.

2.1 Estructuración de los datos

2.1.1 Principales orígenes de información

Dependiendo del nivel de información que aportan las fuentes de información pueden ser tanto primarias o secundarias.

Las ***fuentes primarias*** dominan información de primicia y única, originado de un trabajo intelectual, son considerados como documentos primarios, libros, revistas científicas, documentos oficiales, informes técnicos, entre otros.

Las ***fuentes secundarias*** dominan información fundamentada, transformada, producto de análisis, procedencia o reorganización que representa a documentos primarios originales, son considerados como fuentes secundarias: enciclopedias, libros o artículos que interpretan otros trabajos o investigaciones.

2.1.2 Métodos de recolección de datos

Los métodos de recopilación de datos más conocidos, al momento de obtener la información original de los elementos de análisis que componen la muestra o población por investigar son:

El censo: Es un método de recolección de datos, en la cual la se compone de los individuos que conforman la población o universo bajo estudio.

La encuesta: Es un método de recolección mediante el cual la información se consigue relevando sólo un subconjunto o muestra de elementos del universo en estudio, que permite obtener información sobre el mismo.

La observación: La observación de un fenómeno en estudio es un modo objetivo de recolección. Se puede conseguir información aun cuando no haya el deseo de suministrar y es independiente del contenido y autenticidad de los sujetos a estudiar sin embargo, se debe capacitar al observador, para que la investigación tenga validez científica.

Registro administrativo o archivos: Existen oficinas públicas que llevan registros administrativos para sus propios fines. Por ejemplo, los Registros Civiles, los Ministerios de Educación, las Aduanas, etc.

2.1.3 Agrupamiento de datos

Existen algunos métodos para extractar los datos ya sea medidos u observados. Cuando se trata de variables cualitativas en el que las categorías están establecidas, lo que hay que realizar es la contabilización del número de casos que pertenecen a cada categoría, calculando una proporción, un porcentaje o una razón. Al contrario, cuando se presenta el caso con variables cuantitativas, el extracto de los datos reside en organizar tablas que simplifican los datos originales y se denominan ***distribuciones de frecuencia***.

Frecuencia: Es el número de veces que se ostenta el valor numérico de la variable.

Ejemplo 2.1: Consideremos la observación de 67 escuelas de una localidad urbana, de la cual se cuenta el número de aulas como "variable", resultando que estas varían de 8 a 14. La tabla 2.1 muestra el total de las observaciones.

Tabla 2.1

Datos para generar una distribución de frecuencias

14 13 12 11 10 9 8 9 11 13 14 12 10 8 14 12 10 8 9 11 13 13 11 9 14 11 12 13 8 8 13 14 12 11 9 11 9 13 9 8 10 12 8 10 11 12 13 11 13 12 10 13 12 11 10 12 11 10 12 10 12 12 10 12 12 10 10

Tabla de frecuencias: Representa en forma concreta los distintos valores que pueden representar una variable de acuerdo a sus frecuencias, por ejemplo: De la observación de las 67 escuelas, en las que se ha visto el número de aulas en cada una de ellas, se agrupa las veces que se repite una misma cantidad, desde 8 hasta 14, para tener los datos de forma más resumida.

Tabla 2.2

Agrupación de frecuencias

N° Aulas por Escuela	Frecuencia
8	7
9	7
10	12
11	11
12	15
13	10
14	5
	67

Para definir el rango de valores a utilizar, para la Tabla 2.2 se seleccionaron 7 rangos, se calcula el *intervalo de la clase*.

Intervalo de la clase (*IC*): Es la expresión de los límites que posee la función, con respecto a la dispersión de sus datos.

[2.1] $$IC = \frac{\textit{Valor mas grande} - \textit{Valor mas pequeño}}{\textit{Rango deseado de clase}}$$

Representación gráfica: en lo habitual la representación gráfica de una tabla de frecuencias permite percibir con mayor claridad algunas características de la masa de datos que se investiga. Por ello, a través de gráficos, resulta fácil transmitir conclusiones a personas no habituadas a la interpretación de tablas de frecuencias. Ejemplo: con los datos de la tabla 2.2 se grafica en la ilustración 2.1. Para figurar gráficamente una distribución de frecuencias se manipula un par de ejes de coordenadas.

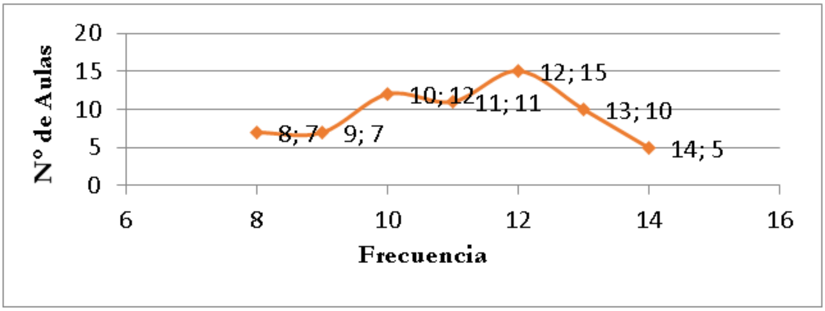


Ilustración 2.1 Representación de la cantidad de aulas en una zona urbana

Parámetros estadísticos: En la ciencia estadística es el indicador que resume una gran cantidad de datos que puede obtenerse del estudio de variables estadísticas.

Como se muestra en la Ilustración 2.1, al conseguir los datos correspondientes a una población, la distribución de frecuencias que se persigue se limita por parámetros, cuyo objetivo es reducir o condensar en pocas cifras el conjunto de observaciones relativas a dichas variables.

2.1.4 Variables según su tipo

En el capítulo 1 se mencionan los tipos de variables que existen al momento de realizar un experimento, recordemos que las variables estadísticas pueden ser tanto cualitativas como cuantitativas.

Variables cualitativas: Son aquéllas que están relacionadas a características de los elementos, estas son de tipo dicotómica o politómica, ordinal o nominal, por lo tanto sus gráficas son representadas en gran parte de forma sectorial.

Variable cualitativa dicotómica: Cuando sólo alcanzan escoger dos opciones posibles como sí y no, hombre y mujer o son politómicas cuando alcanzan adquirir tres o más opciones de valores a escoger.

Variable cualitativa ordinal: La variable consigue brindar distintos valores ordenados alcanzando una escala establecida, aunque no es obligatorio que el intervalo entre exactitudes sea uniforme, por ejemplo: grave, moderado, leve.

Variables cualitativas nominal: En esta variable los valores no consiguen ser sometidos a una razón de orden como por ejemplo: el lugar de residencia o los colores.

Variables cuantitativas o métricas: Son las variables numéricas, que se pueden cuantificar o medir, se clasifican en dos tipos: contínuas o discretas.

Variable cuantitativa discreta: Esta variable presenta alejamientos o interrupciones en la escala de valores que puede escoger. Estas dispersiones o interrupciones indican la ausencia de valores entre los diferentes valores determinados que la variable pueda adjudicarse. Un ejemplo, es el número de hijos en un hogar.

Variable cuantitativa contínua: Esta variable puede conseguir cualquier valor entre de un intervalo especificado de valores. Por ejemplo la altura o el peso, que simplemente es restringido por la precisión del aparato medidor, en el supuesto permiten que constantemente exista un valor entre dos cualesquiera.

Tabla 2.3

Clasificación de las variables estadísticas

Variable	Tipo de variable	Atributos	Ejemplo
Variable Cualitativa	Cualitativa Dicotómica	Solo tiene dos opciones posibles	Si, No; Masculino, Femenino
	Cualitativa Ordinal	Expresa distintos valores en una escala	Alto, medio, bajo; grave, moderado, leve
	Cualitativa Nominal	Los valores no tienen orden	Belleza, miedo, hambre, conocimiento
Variable Cuantitativa	Cuantitativa Discreta	Expresan unidades de tipo cuantificable	Número de integrantes de una familia, pacientes de un hospital, Empleados de una fábrica
	Cuantitativa Contínua	Expresan unidades de tipo medible	Peso, altura, densidad, temperatura

2.2 Tipos de gráficas para distribuciones de variables cualitativas

2.2.1 Diagramas de barras

Es uno de los métodos gráficos más representativos, debido por su sencillez, para presentar las características cualitativas, aún en aquellos casos en que la búsqueda de información no está dada en tablas de frecuencias.

Ilustración 2.2 Diagrama de barras

2.2.2 Diagrama circular o de sectores

Se recurre a este método gráfico con mucha frecuencia para personificar características cualitativas y se usa para destacar las diferencias en las proporciones o porcentajes en que está dada la distribución.

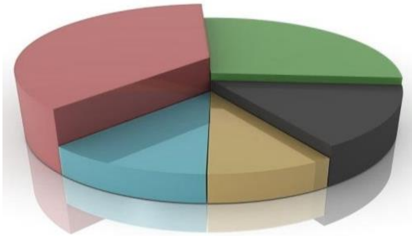


Ilustración 2.3 Diagrama circular o de sectores

2.2.3 Pictograma

Es una manera de personificar las cantidades estadísticas por medio de dibujos, manipulando para ello objetos y figuras.

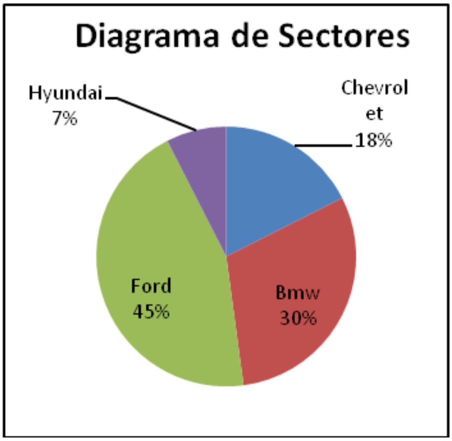
Ilustración 2.4 Pictograma
Fuente: es.paperblog.com/pictogramas-y-cartogramas

Ejemplo 2.2: Realiza los diferentes tipos de gráficas para los siguientes datos:
La empresa automotriz Ford desea saber cual es su participación de mercado, con respecto a la competencia, para esto realiza un estudio en el cuál obtiene la información de ventas anuales de su competencia, que es la siguiente: Chevrolet: 124 autos; BMW: 214 y Hyundai: 53.

Tabla 2.4

Mercado automotriz

Marca de Auto	Chevrolet	BMW	Ford	Hyundai	Total
Ventas	124	214	315	53	706
Porcentaje	0,18	0,3	0,45	0,08	1



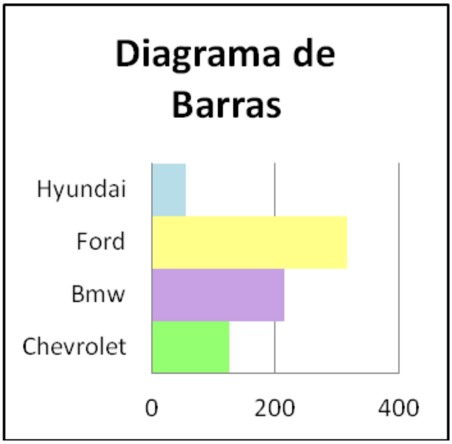
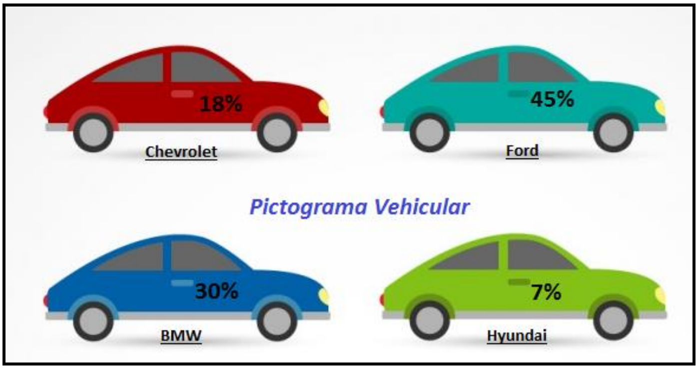


Ilustración 2.5 Desarrollo de gráficos



2.3 Representación gráfica de tablas de distribución de frecuencias para variables discretas

2.3.1 Gráficos para variables discretas

Cuando se desea representar una variable discreta, usamos el diagrama de barras, con esto pretendemos hacer una gráfica diferencial donde las variables son comparadas unas con otras. Las barras deben ser estrechas para representar los valores que toma la variable discreta. El diagrama de las variables hace que sea en forma de escalera.

Ejemplo 2.3: Se considera realizar el estudio de un curso con niños de pre-escolar con edades de entre 3 a 6 años.

5	6	5	4	3
6	3	4	5	4
3	4	6	5	3
4	3	6	4	6

Para representar gráficamente el resultado, en primer lugar observamos que la variable X es de tipo cuantitativa discreta, porque su resultado esperado es de carácter numérico, lo que expresamos de la siguiente manera:

$$X_i \in 3, 4, 5,6$$

A continuación se ordenan los resultados estadísticamente.

Tabla 2.5

Tabla de distribución de frecuencias de la edad de los niños

Edad (variable xi)	Nº de niños (Frecuencia)	Frecuencia Acumulada
3	5	5
4	6	11
5	4	15
6	5	20
Total	20	

Ilustración: Diagrama de barras es completo cuando es para una variable discreta.

Cumpliendo que el diagrama integral (progresivo) contabiliza el número de observaciones de la variable inferior o igual a cada punto del eje de abscisas.

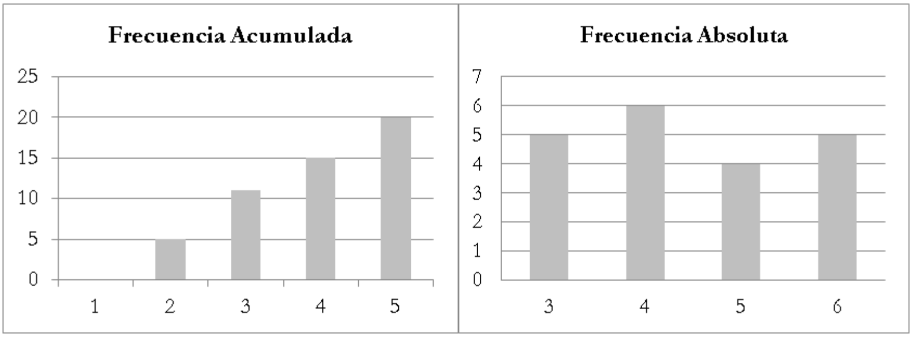


Ilustración 2.6 Frecuencias absolutas y acumuladas

Ejemplo 2.3: Al momento de un censo se encuentran clasificadas 12 familias por su número de hijos y se obtuvo el siguiente cuadro de distribución. Comparar los diagramas de barras para frecuencias absolutas y relativas. Realizar el diagrama acumulativo creciente.

Número de hijos (xi)	1	2	3	4
Frecuencias (ni)	1	3	5	3

En primer lugar, se registra la tabla de frecuencias en el modo habitual:

Tabla 2.6

Número de miembros del hogar

Variable	Frecuencia	Frecuencia	Frecuencia
	Absoluta	Relativa	Acumulada
xi	**Ni**	**fi**	**Ni**
1	1	0,083	1
2	3	0,25	4
3	5	0,416	9
4	3	0,25	12
	12	1	

Con las columnas relativas a x_i y n_i se realiza el diagrama de barras para frecuencias absolutas, lo que se muestra en la figura siguiente.

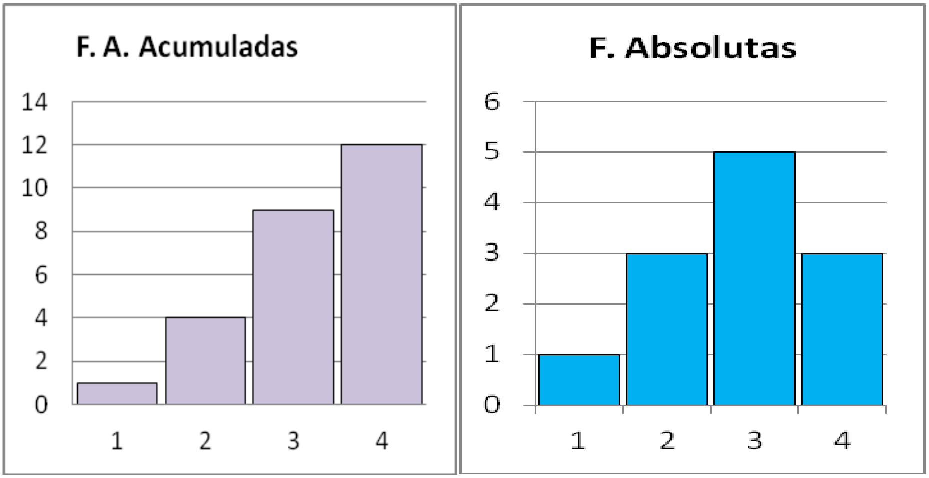


Ilustración 2.7 Frecuencias absolutas y acumuladas

2.4 Elaboración y gráficas de tablas de distribución de frecuencias para variables contínuas

Cuando las variables son contínuas, se utiliza en la representación de los valores, los diagramas diferenciales, histogramas y los polígonos de frecuencias.

Un ***histograma*** se desarrolla a partir de la tabla estadística, incorporando sobre cada intervalo, un cuadrilátero que posee este segmento como base. La razón para calcular la altura de cada cuadrilátero es conservar la proporción entre las frecuencias absolutas o relativas que posee en cada intervalo y el área.

El ***polígono de frecuencias*** se diseña simplemente si tenemos representado anticipadamente el histograma, ya que radica en unir mediante líneas rectas los topes del histograma que pertenecen a las marcas de clase. Hay que observar, que de este método, un polígono de frecuencias posee características en común con el histograma, ya que las áreas de las gráficas son idénticas.

El ***diagrama integral*** también conocido como polígono de frecuencias acumulado, se obtiene como la poligonal delimitada en las abscisas, desde los extremos de los intervalos en los que tenemos organizada la tabla de la variable, y en las ordenadas, las alturas que son proporcionales a las frecuencias acumuladas.
Ejemplo 2.4: Con la siguiente distribución de variables cuantitativas contínuas, elaborar las gráficas.

Tabla 2.7
Distribución cuantitativa contínua

Límites	Marca de Clase ci	Frecuencia Absoluta fi	Frecuencia Acumulada Fi	Frecuencia Relativa ni	Relativa Acumulada Ni
[0, 5)	2.5	1	1	0.025	0.025
[5, 10)	7.5	1	2	0.025	0.050
[10, 15)	12.5	3	5	0.075	0.125
[15, 20)	17.5	3	8	0.075	0.200
[20, 25)	22.5	3	11	0.075	0.275
[25, 30)	27.5	6	17	0.150	0.425
[30, 35)	32.5	7	24	0.175	0.600
[35, 40)	37.5	10	34	0.250	0.850
[40, 45)	42.5	4	38	0.100	0.950
[45, 50)	47.5	2	40	0.050	1
		40		1	

Para calcular la marca de clase (Ci) se elabora un promedio para cada intervalo, como referencia en la tabla 2.7 se toma la segunda fila.

$$C_i = \frac{10 + 5}{2} = 7.5$$

El cálculo de las frecuencias acumuladas es la suma del número de frecuencia actual con el anterior, en el área sombreada se suma la frecuencia de los intervalos [0,5) y [5,10), dando como respuesta 2 frecuencias.

Para el análisis de las frecuencias relativas solo es la división entre el número de frecuencia de los intervalos y la sumatoria de éstos, generando un porcentaje representativo de cada intervalo.

$$n_i = \frac{1}{40} = 0.025$$

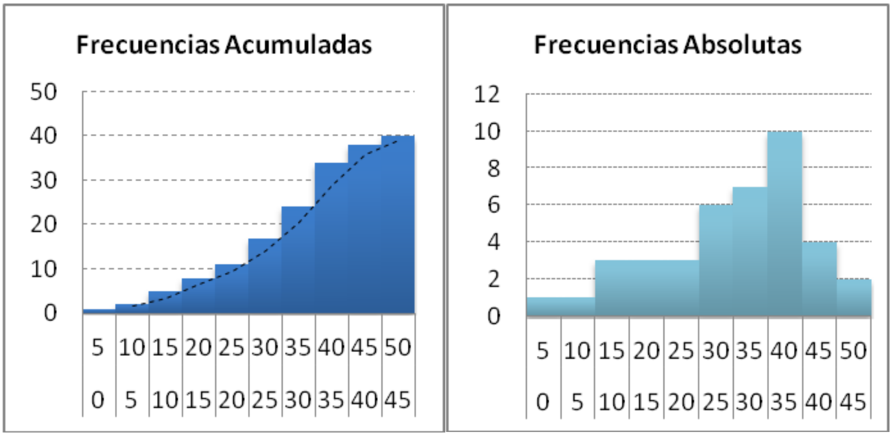


Ilustración 2.5 **Gráfico de frecuencias absolutas y acumuladas**

Ejemplo 2.5: La siguiente distribución se trata de una duración en horas (completas) de un lote de 500 latas, pertenecientes a una empresa siderúrgica:

Tabla 2.8

Distribución de empresa siderúrgica

Límites	**fi**	**Frecuencia Absoluta fi**	**Frecuencia Relativa ni**	**Relativa Acumulada Ni**	**Variación de 900 horas**
300-500	200	50	0,1	0,1	0,1
500-700	200	150	0,3	0,4	0,3
700-1.100	400	275	0,55	0,95	**0,275**
1.100-1.300	200	25	0,05	1	0,05
		N=500	1		

- Constituir el histograma de frecuencias relativas y el polígono de frecuencias.
- Diseñar la curva de frecuencias relativas acumuladas.
- Establecer el número mínimo de latas que tienen una duración inferior a 900 horas.

a) Histograma de frecuencias

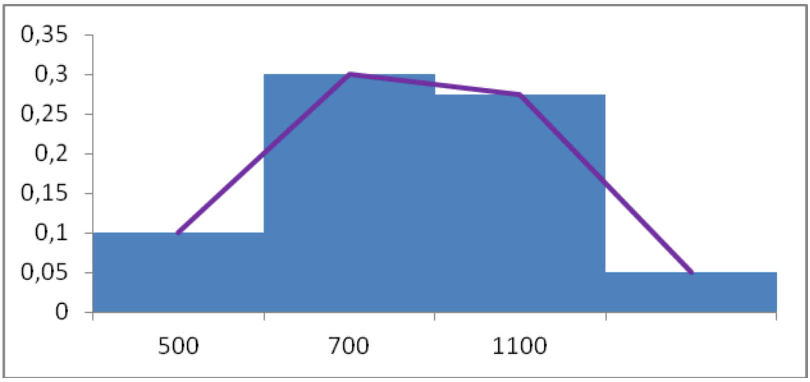


Ilustración 2.6 Resultado de frecuencias absolutas

b) Curva de frecuencias

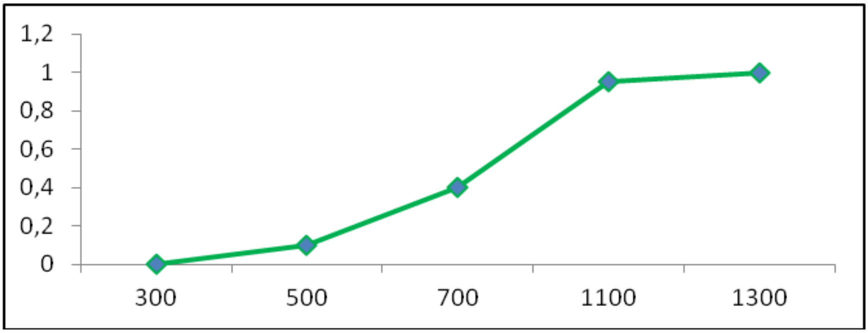


Ilustración 2.7 Resultados de frecuencias acumuladas

c) Por otro lado, se observa que la figura b se ve que sumando frecuencias relativas, hasta las 900 horas de duración hay 67,5 % de las latas:

$$0{,}10 + 0{,}30 + 0{,}275 = 0{,}675$$

Como en total son 500 latas, el número de latas con una duración tan igual o menor que 900 horas es redondeando, 338 latas.

$$0{,}675 * 500 = 337{,}5 \text{ equivalente a } 338 \text{ latas.}$$

2.5 Ejercicios propuestos

1- Supongamos que el supervisor de ventas averigua los precios (en $) de cierto artículo en 40 almacenes diferentes y encuentra los siguientes datos:

76	85	80	88	74	65	71	83
76	83	71	70	86	67	68	73
77	71	75	75	68	74	72	75
84	75	75	73	87	68	72	70
72	63	77	88	60	72	83	88

Se pide elaborar una tabla de frecuencias para esta variable discreta y las gráficas correspondientes.

2.- Se ejecuta una observación en el barrio a 150 hogares de clase media, para saber la clase de proteína o alimento se consumía en la cocina, los resultados son a continuación: res, 14 hogares; pollo, 65 hogares; pescado, 21 hogares; compran embutidos sin especificaciones de clase, 17 hogares; huevos, 21 hogares; cerdo, 6 hogares; otras aves, 13 hogares.

- ¿Cuál es la población?
- ¿Cuál es el tamaño de la muestra?
- Construir una distribución de frecuencias en el mismo orden indicado.
- ¿Cómo se explica que la suma de las frecuencias sea superior al número de hogares?

3.- La fábrica de gaseosas Coca-Cola proyecta lanzar al mercado un nuevo sabor de bebida baja en azúcar. Se realiza un estudio de aprobación de dicho sabor en una muestra de 30 niños, manejando una escala de 10 puntos, para medir el grado de aceptación. Los puntos conseguidos en los 30 niños fueron los siguientes:

2 6 8 7 4 5 10 6 6 7 6 7 3 8 7 6 8 6 5 4 7 8 5 7 6 7 2 7 2 7

La muestra estuvo formada por igual número de individuos de ambos sexos, de 6 a 12 años, residentes a una institución escolar del barrio El Edén de la ciudad de Cuenca. Elabore las tablas y gráficas correspondientes.
4.- Realiza un diagrama de sectores y de barras para los siguientes datos:
Servicios Médicos Públicos Nº de Centros de Salud
Guayaquil 30
Durán 27
Samborondón 43
Milagro 25
Daule 40
Nobol 15

5.- La representación de veces que han ido al cine en el último mes un grupo de amigos es:
X = {2,3,0,1,5,3,2,1,0,0,2,1,2,3,5,0,5,4,1,1,1,2,0,1,2}

Desarrollar la tabla de frecuencias absolutas y relativas, y las acumuladas.

6.- Se efectúa un trabajo grupal en la asignatura de estadística en una clase formada por 40 estudiantes. 2 estudiantes realizan el trabajo en una, 5 en 2 carpetas, 6 en 3 carpetas, y el resto en 4 carpetas. Forma la tabla de frecuencias con los datos gráficos ¿Crees que el profesor ha recomendado un número determinado de carpetas?

7.- Realiza un diagrama de sectores para los siguientes datos:
Color del pelo Nº
Moreno 124
Castaño 214
Rubio 315
Pelirrojo 53

8.- En una universidad se ha elaborado una encuesta a 200 estudiantes de la carrera de Ingeniería Civil. El 32% alega que está muy contento con la universidad, el 40% está contento, el 23% no está contento y el resto muy descontento. Forma la tabla de frecuencias absolutas, relativas, absolutas y las gráficas.

9.- Completa la siguiente la tabla de frecuencias absolutas y relativas.

Clase	Frecuencia absoluta	Frecuencia relativa
[0 - 10)		
[10-20)	7	0,35
[20-30)		0
[30-40)	2	0,1
[40-50)	8	
Total	20	1

10.- El número de veces que han ido a una conferencia durante el último mes los estudiantes universitarios es:

Nº veces fi

0 15

1 26

2 32

3 20

4 15

5 o más 8

Realiza el polígono de frecuencias para estos datos.

11.- Los aviones que han arribado en el aeropuerto José Joaquín de Olmedo lo hicieron en los subsiguientes minutos de cada hora:

{0,13,23,55,57,43,32,23,47,50,0,12,14,27,34,56,3,12,34,39,46,10,15,30,45,0,24,37,58}

Estructurar la tabla de frecuencias absolutas y relativas, y acumuladas, acomodándolos en clase por cuarto de hora.

12.- En una productora avícola se han tomado los pesos (en kg) de 50 pavos:
2,8 - 3,2 - 3,8 - 2,5 - 2,7 - 3,7 - 1,9 - 2,6 - 3,5 - 2,3 - 3,0 - 2,6 - 1,8 - 3,3 - 2,9 - 2,1 - 3,4 - 2,8 - 3,1 - 3,9 2,9 - 3,5 - 3,0 - 3,1 - 2,2 - 3,4 - 2,5 - 1,9 - 3,0 - 2,9 - 2,4 - 3,4 - 2,0 - 2,6 - 3,1 - 2,3 - 3,5 - 2,9 - 3,0 - 2,7 2,9 - 2,8 - 2,7 - 3,1 - 3,0 - 3,1 - 2,8 - 2,6 - 2,9 - 3,3

a) Construye una tabla con los datos agrupa dos en 6 intervalos de amplitud 0,4 kg.
b) Representa gráficamente esta distribución.

2.6 Ecuaciones introducidas en el capítulo 2

Intervalo de la clase (*IC*): es la expresión de los límites que posee la función, con respecto a la dispersión de sus datos.

[2.1]
$$\boldsymbol{IC = \frac{Valor\ mas\ grande - Valor\ mas\ peque\tilde{n}o}{Rango\ deseado\ de\ clase}}$$

Unidad 3

ESTADÍSTICA SUMARIA, PRINCIPALESMEDIDAS Y SUS APLICACIONES

"El éxito moderado se puede explicar por las habilidades y el trabajo. Un éxito enorme sólo es atribuible a la varianza."

-NASIM TALEB-

Escritor, ensayista, profesor, filósofo, economista e investigador estadounidense nacido en el Líbano, autor de "¿Existe la suerte?" (2001), "El cisne negro: el impacto de lo altamente improbable" (2007), "El lecho de Procusto" (2010) y "Antifrágil: las cosas que se benefician del desorden" (2012).

CONTENIDO DE LA UNIDAD

INTRODUCCIÓN

El desarrollo del siguiente capítulo tiene como propósito, estudiar los métodos empleados para resumir los datos originados dentro de un estudio, en indicadores que sinteticen su estructura y comportamiento por ejemplo los promedios, dispersión, variabilidad, asimetría, etc. Actualmente se incorporan estos métodos para estudiar o medir el comportamiento de los elementos que constituyen una población, si bien es cierto que los cuadros o gráficas, describen un fenómeno, no lo hacen en forma completa; por tanto se requiere acudir a ciertas medidas denominadas parámetros o valores estadísticos de la población.

OBJETIVOS DE APRENDIZAJE

Al finalizar el capítulo, será capaz de:

- Desarrollar destrezas en el estudio de las distintas medidas de tendencia central.
- Interpretar los resultados logrados mediante la aplicación de promedios.
- Diferenciar cuál es el promedio que debe ser utilizado según las circunstancias.
- Calcular los parámetros estadísticos y explicar las características, usos, ventajas y desventajas de cada medida de dispersión.
- Interpretar el sinificado de las medidas de dispersión

Las medidas de tendencia cental y de dispersión son importantes porque ayudan a la toma de decisiones dentro de una empresa, a través de indicadores como, resultados de balances, costos, gastos generales, volumen de producción, inventarios, ventas, cuentas de resultados y otros aspectos que atañen a la actividad interna, y de aquellos que provienen de fuera de la empresa, la marcha de la actividad industrial o comercial.

3.1 La estadística sumaria y su importancia

El capítulo inicia con el estudio de la estadística descriptiva, para transformar un grupo de datos en bruto, en algo considerado de mayor importancia, algún método o medida más exacto. En estos casos, se puede usar los métodos que constituyen la estadística sumaria para describir las características principales de un conjunto de datos.

Existen cuatro características que son de mucha importancia en particular, las cuales sirven para el análisis de una variable o de una distribución unidimensional:

- Medidas de tendencia central.
- Medidas de posición no centrales.
- Medidas de dispersión o de variabilidad.
- Medidas de asimetría o formación.

3.2 Medidas de tendencia central

Son empleadas para describir o sintetizar mediante un número único, designado como promedio, esto expresa que es la posición de un valor en la variable, de tal modo que represente al grupo de valores observados, un promedio es un resultado que intenta representar o simplificar las características notables de un grupo de valores.

El promedio es un término genérico y es aceptable su función mientras no se reseñe a una de las medidas de posición en forma determinada.

Los promedios reciben el nombre genérico de medidas de tendencia central porque algunos constituyen valores ubicados en el centro de la variable a la cual representan. Se consideran varias clases de promedios o medidas de posición:

- Media aritmética.
- Media geométrica.
- Media armónica.
- Mediana.
- Moda.
- Percentiles y cuartiles.

3.2.1 Media aritmética

Es la medida más conocida debido a su forma sencilla de calcular y por supuesto con la que siempre estamos familiarizados, ya que siempre se ha calculado el promedio de calificaciones obtenida en un período escolar determinado. Ocasionalmente, se le denomina meramente media o promedio y es usado con mucha frecuencia.

La media aritmética presenta algunas ventajas: es el único promedio que se presta a tratamientos algebraicos, muestra una estabilidad en el muestreo y es altamente sensible a cualquier variación en los valores de la distribución. Su mayor desventaja reside en la imposibilidad de ser empleada en aquellas distribuciones que no poseen definidos sus elementos extremos y debido a su perceptibilidad para elementos grandes de la variable, consigue dar un valor promedio que no sea característico o típico. Debido a que la media aritmética se ve influida por la existencia de valores extremos que pueden dar lugar a existencia de un promedio que no sea representativo de la serie.

3.2.2 Media aritmética (simple)

La media aritmética se define como el cociente que se obtiene al dividir la suma de los valores de la variable por el número total de observaciones.

Una muestra con n observaciones, tiene una media χ denominada media estadística, del elemento de una muestra.

[3.1] $$\boldsymbol{x} = \frac{\sum \boldsymbol{x}}{\boldsymbol{n}}$$

Una población con N elementos tiene una media μ denominada parámetro de los datos de una población completa

[3.2] $$\mu = \frac{\sum \boldsymbol{X}}{\boldsymbol{N}}$$

Ejemplo 3.1: El Ministerio de Inclusión Social ofrece un estímulo especial a aquellas fundaciones en las que la edad promedio de los niños que asisten está por debajo de 9 años. Si los siguientes datos corresponden a las edades de los niños que acuden de manera regular a los Centros ¿calificará éste para el estímulo?

Datos de un centro de educación: 8, 5, 9, 10, 9, 12, 7, 12, 13, 7, 8.

$$x = \frac{\sum x}{n} = \frac{8+5+9+10+9+12+7+12+13+7+8}{11} = 9{,}09$$

La media muestral corresponderá a una edad promedio de 9,09 años

Ejemplo 3.2: Calcular el promedio de las calificaciones de obtenidas por los alumnos que pertenece a la materia de Estadística del tercer año de universidad del paralelo 2ª

X = 5, 10, 16, 14, 12, 17, 16.

Si el paralelo 2A es el único disponible, nos encontramos en una media de tipo poblacional. Por lo que su resultado se obtendría:

$$\mu = \frac{\sum X}{N} = \frac{5+10+16+14+12+17+16}{7} = 12{,}86$$

El promedio de las calificaciones corresponderá a 12,86

3.2.3 Media aritmética ponderada (datos agrupados)

Cuando el número de observaciones es grande, las operaciones para calcular la media aritmética se simplifican si agrupamos los datos en una tabla de frecuencias. Para calcular la media ponderada, primero calculamos el punto medio de cada elemento, después multiplicamos cada punto medio por la frecuencia absoluta de cada intervalo.

[3.3] $$\boldsymbol{x = \frac{\sum m_i f_i}{n}}$$

Donde las variables

$\boldsymbol{m_i}$ **:** Punto medio de la variable.

$\boldsymbol{f_i}$: Frecuencia de la variable.

Ejemplo 3.3: Con la siguiente tabla de frecuencias, calcule la media ponderada.

Tabla 3.1

Frecuencias con intervalos

INTERVALO	***mi***	***Fi***
[50-60)	55	10
[60-70)	65	18
[70-80)	75	14
[80-90)	85	6
[90-100)	95	2

Se calcula el valor de las frecuencias de la tabla.

Tabla 3.2

Cálculo de la media ponderada

INTERVALO	***Mi***	***fi***	***mi*fi***
[50-60)	55	10	550
[60-70)	65	18	1170
[70-80)	75	14	1050
[80-90)	85	6	510
[90-100)	95	2	190
		50	**3470**

$$x = \frac{\sum m_i f_i}{n} = \frac{3470}{50} = 69{,}4$$

3.2.4 Mediana (Me)

Se define como aquel valor de la variable que destaca la mitad de las observaciones y ocupa el valor central de estas.

Por esta razón, se le supone como el valor central, ya que el promedio está situado en el centro de la distribución. Su estudio es menos frecuente que la media aritmética, presenta inestabilidad en el muestreo, sus fórmulas son rígidas y no admiten procedimiento algebraico como la media. En las distribuciones irregulares, que presentan valores extremos que por lo general afectan al promedio, debe usarse la mediana, ya que no se afecta por las variaciones que sufra la variable, mientras no sea en la observación central. Para realizar el cálculo de la mediana se requiere una clasificación de los datos, de menor a mayor o viceversa.

Datos no agrupados

Se presentan dos situaciones:

Número impar de datos: La mediana es el dato que está en la posición $\frac{n+1}{2}$

[3.4] $$Me = X_{(\frac{n+1}{2})}$$

Ejemplo 3.4: Sea el conjunto ordenado de datos:

2	3	5	6	8
x1	x2	x3	x4	x5

$$Me = X_{(\frac{5+1}{2})} = X_{(3)} = 5$$

Número par de datos: Es el promedio entre los dos datos centrales.

[3.5] $$Me = \frac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2}$$

Ejemplo 3.5: Sea el conjunto ordenado de datos:

2	3	5	6	8	9
x1	x2	x3	x4	x5	x6

$$Me = \frac{X_{(3)} + X_{(4)}}{2} = \frac{5+6}{2} = 5,5$$

Datos agrupados

Si los datos están concentrados en una distribución de frecuencias, se elige el intervalo de clase que domina a la mediana conocido como clase mediana. Para ello, debemos establecer la frecuencia acumulada absoluta que posee al elemento número $\frac{n+1}{2}$ **.** El valor de este intervalo para la mediana se calcula usando la siguiente ecuación.

[3.6] $$Me = X = L_m + \left[\frac{\frac{n+1}{2} - F_{i-1}}{f_i}\right] * h$$

Dónde:

- Me: Mediana.
- Lm: Límite inferior de clase mediana.
- n: Número de datos.
- Fi-1: Frecuencia acumulada absoluta de la clase anterior al intervalo de la mediana.
- Fi: Frecuencia absoluta de la clase mediana.
- h: Amplitud del intervalo.

Ejemplo 3.6: Las edades de las personas que pertenecen a una casa de retiro está representada por la siguiente tabla:

Tabla 3.3

Edades de personas en la casa de retiro

EDAD	*Me*	*fi*	*fri*	*fri%*	*Fi*	*Fri*	*Fri%*
[50-60)	55	10	0,2	20	10	0,26	20
[60-70)	65	18	0,4	36	28	0,566	56
[70-80)	75	14	0,3	28	42	0,84	84
[80-90)	85	6	0,1	12	48	0,96	96
[90-100)	95	2	0	4	50	1	100

La clase mediana es la que contenga el elemento en la posición $\frac{50+1}{2}$, es decir en la posición **25,5**. Buscamos en la frecuencia acumulada **Fi** y vemos que se halla en el intervalo **[60, 70)**

$$Me = 60 + \left[\frac{25,5 - 10}{18}\right] * 10 = 68,61$$

3.2.5 La moda (Mo)

Es el valor que más se repite en un grupo de datos. La moda es otra medida de posición y su uso es bastante limitado. Al igual que la mediana, sus fórmulas no aceptan tratamiento algebraico tampoco es sensitivo a valores extremos o a la variación que se hagan a los valores de la variable diferentes al de la moda. Su utilización se hace necesaria cuando la distribución presenta el primero y último intervalo abierto o no definido.

Ejemplo 3.7: Inspeccionemos tres casos de observaciones para situar en ellos en el valor de la moda:

a) 4, 9, 10, 11, 15, 13, 14, 15;	Moda (15) por que se repite dos veces
b) 3, 4, 6, 8, 10, 12, 13, 15;	Sin Moda, no hay elemento que se repita
c) 4, 4, 5, 8, 10, 4, 11, 13, 10;	Bimodal (4 y 10)

3.2.6 Media geométrica o cuadrática (Mg)

La media geométrica es un promedio que además se encuentra definido rigurosamente por una fórmula matemática, ver [3.7] esta se utiliza cuando se pretende dar importancia a valores pequeños de la variable o cuando se quiere obtener el promedio de una sucesión de valores que están dados en progresión geométrica o aproximadamente geométrica. Su función en el campo industrial y comercial es muy restringida y su utilidad se delimita a la preparación de promedios sobre el crecimiento o decrecimiento en una variable.

Datos no agrupados u originales

El promedio geométrico, se calcula hallando el producto de todos los elementos de la serie y luego extrayendo la raíz del orden del número de observaciones consideradas.

[3.7] $$Mg = \sqrt[n]{\prod_{i=1}^{n} x_i}$$

Ejemplo 3.8: Un país tiene en el año 2010 una población de 14 millones, la que sube a 171,380 millones en el 2011. Se pregunta por la población media en dicho período.

$$Mg = \sqrt[2]{14(171,380)} = 1,549$$

Hay un total de 1, 549,000 habitantes

Datos agrupados

La media geométrica se define como la raíz enésima del producto de los valores de la variable, elevadas cada una de ellas a una potencia, la cual está dada por la frecuencia absoluta.

[3.8] $$\boldsymbol{Mg = \sqrt[n]{\prod_{i=1}^{n} y_i^n}}$$

Al igual que en datos no agrupados, las operaciones se simplifican si trabajamos con logaritmos:

[3.9] $$\boldsymbol{LogMg = \frac{\sum n_i logy}{n}} \ ; \boldsymbol{Mg = \ anti\ log \frac{\sum n_i logy}{n}}$$

Ejemplo 3.9: Consideremos los datos de la siguiente tabla para calcular la media geométrica.

Tabla 3.4

Cálculo de la media geométrica

Yi-1 - Yi	***yi***	***Ni***	**log *y***	***n* log *y***
46,1 - 54	50	3	1,6989	5,09691
54,1 - 6	58	6	1, 6343	10,58058
6,1 - 0	66	10	1,81954	18,1954
0,1 - 8	4	6	1,869 3	11,1538
8,1 - 86	8	3	1,91381	5, 4143
86,1 - 94	90		1,954 4	3,90848
		30		54, 3818
Li - Ls	*i*	*Fi*	log *i*	*fi** log *i*

$$\boldsymbol{Mg} = \textbf{antilog}\left(\frac{\mathbf{54,3818}}{\mathbf{30}}\right) = \boldsymbol{antilog}(\mathbf{1,812726}) = \ \mathbf{10^{1,812726}} = \mathbf{64,97}$$

3.2.7 Media armónica (H)

El promedio armónico se simboliza por H. Este promedio se precisa expresando que el recíproco de la media armónica es igual a la media aritmética del recíproco de los valores de la variable. Los elementos del conjunto deben ser precisamente

no nulos. Esta media es poco sensible a los valores grandes, pero muy sensible a los valores próximos a cero, ya que los recíprocos 1/Xi son muy altos.

[3.10] $$\boldsymbol{H = \frac{n}{\sum_{i=1}^{n} \frac{1}{Xi}}}$$

Ejemplo 3.10: Un camión realiza un recorrido de 400km. La carretera en mal estado que no permitían correr al comienzo de 100 km los recorre a 120km/h; los consecutivos 100km la vía está en demasiado mal estado y va a 20km/h; los terceros los recorre a 100km/h y los 100 últimos a 130km/h. Para calcular el promedio de velocidades, calculamos la media armónica.

$$\boldsymbol{H = \frac{4}{\frac{1}{120} + \frac{1}{20} + \frac{1}{100} + \frac{1}{130}} = \frac{4}{\frac{593}{7,800}} = \frac{31,200}{593} = 52,61}$$

3.2.8 Diferencia entre los promedios

La Media Aritmética se caracteriza porque se ve influida por la existencia de valores extremos en la serie, es decir si hay un dato muy alejado de los otros, tanto hacia arriba o hacia abajo, la media aritmética se verá afectado por ese valor.

Un ejemplo claro es el cálculo de los saldos promedio de las cuentas bancaria, donde se toman todos los saldos al cierre diario, acumulados todos los días, para calcular el promedio del mes.

Si durante los 30 días se ha mantenido un saldo homogéneo, el promedio aritmético simple resultante será un valor representativo de la serie, pero si en un día determinado, un cliente tiene un sobregiro en la cuenta, es decir un saldo negativo, el promedio del mes tenderá a bajar.

Por el contrario si dentro del mes, el cliente realizar un depósito muy considerable, este saldo promedio aritmético simple tenderá a subir, por lo cual la medida adecuada en estos casos, en que la variabilidad es muy alta, es la Mediana, puesto que la Mediana no está influida por los valores muy altos o muy bajos.

La Mediana es el valor central dentro de la serie, por lo tanto es muy aconsejable su uso para el caso de datos muy dispersos, por ejemplo aquellos en que se denota diferencias significativas como por ejemplo: los ingresos promedio de una comunidad, o los salarios promedio en una empresa, donde existen niveles gerenciales que tienen un nivel de sueldo alto, en detrimento de otro grupo que tiene un ingreso menor.

La Moda es una medida de tendencia central que se utiliza básicamente cuando se requiere calcular un promedio para datos cualitativos, como por ejemplo, el equipo de fútbol que en promedio es preferido por un grupo de personas, a las cuales se les ha realizado dicha pregunta en una encuesta, el número de zapato promedio que se vende en un almacén, el plato de comida preferido en una localidad, etc.

- Si la media es igual moda y esta igual a la mediana, la distribución es simétrica.
- Cuando la media es mayor que la mediana, la distribución es asimétrica con una cola a la derecha (sesgada a la derecha).
- Cuando la media es menor que la mediana, la distribución es asimétrica con una cola a la izquierda (sesgada a la izquierda).

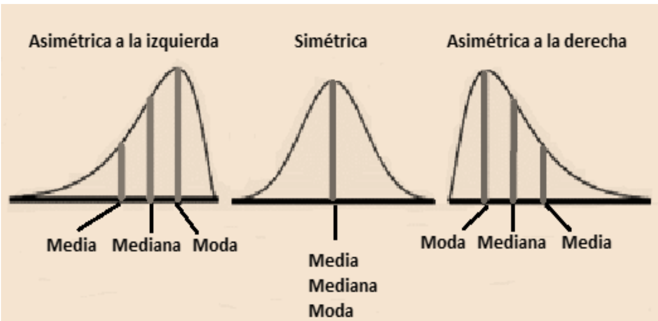


Ilustración 3.1 Diferencia gráfica de las medidas de tendencia central

3.3 Medidas de posición no centrales

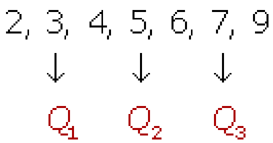
3.3.1 Cuartiles

Los cuartiles son los valores que dividen un conjunto de datos ordenados en cuatro partes porcentualmente iguales. Los cuartiles Q_1, Q_2 y Q_3 establecen los valores proporcionados al 25%, al 50% y al 75% de los datos. Además el Q_2 debe coincidir con la mediana.

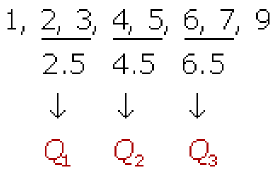
Cálculo de los cuartiles

1. Se Ordenan los datos de menor a mayor.
2. Luego se busca el lugar que ocupa cada **cuartil** mediante la expresión $\frac{\boldsymbol{k*N}}{\boldsymbol{4}}; \boldsymbol{k = 1, 2, 3, 4}$.

Número impar de datos X = {2, 5, 3, 6, 7, 4, 9}



Número par de datos X = {2, 5, 3, 4, 6, 7, 1, 9}



Ejemplo 3.11: Calcular los cuartiles de la distribución de la tabla:

Tabla 3.5

Cálculo de los Cuartiles

Intervalos	f_i	F_i
[50, 60)	8	8
[60, 70)	10	18
[70, 80)	16	34
[80, 90)	14	48
[90, 100)	10	58
[100, 110)	5	63
[110, 120)	2	65
	65	

Primer cuartil:

$$\frac{\mathbf{65*1}}{\mathbf{4}} = \mathbf{16,25}\ ;\ Q_1 = 60 + \frac{16{,}25-8}{10} * 10 = 68{,}25$$

Segundo cuartil:

$$\frac{\mathbf{65*2}}{\mathbf{4}} = \mathbf{32,5}\ ;\ Q_2 = 70 + \frac{32{,}5-18}{16} * 10 = 79{,}06$$

Tercer cuartil:

$$\frac{\mathbf{65*3}}{\mathbf{4}} = \mathbf{48,75}\ ;\ Q_3 = 90 + \frac{48{,}75-48}{10} * 10 = 90{,}75$$

3.3.2 Deciles

Los deciles son los nueve valores que fragmentan la serie de datos en diez partes iguales. Los deciles dan los valores proporcionados al 10%, al 20%... y al 90% de los datos. El decil D5 debe coincidir con la mediana.

Cálculo de los deciles

En parte averiguamos la clase donde se encuentra $\frac{\boldsymbol{k*N}}{\mathbf{10}}$**;** en la tabla de las frecuencias acumuladas.

Ejemplo 3.12: Calcular los deciles de la distribución de la tabla:

Tabla 3.6

Cálculo de los decíles

Intervalos	f_i	F_i
[50, 60)	8	8
[60, 70)	10	18
[70, 80)	16	34
[80, 90)	14	48
[90, 100)	10	58
[100, 110)	5	63
[110, 120)	2	65
	65	

Primer decil:

$$\frac{\mathbf{65*1}}{\mathbf{10}} = \mathbf{6,5}\ ; D_1 = 50 + \frac{6{,}5-0}{8} * 10 = 58{,}12$$

Segundo decil:

$$\frac{\mathbf{65*2}}{\mathbf{10}} = \mathbf{13}\ ; D_2 = 60 + \frac{13-8}{10} * 10 = 65$$

Tercer decil:

$$\frac{\mathbf{65*3}}{\mathbf{10}} = \mathbf{19,5}\ ; D_3 = 70 + \frac{19{,}5-18}{16} * 10 = 70{,}94$$

Cuarto decil:

$$\frac{\mathbf{65*4}}{\mathbf{10}} = \mathbf{26}\ ; D_4 = 70 + \frac{26-18}{16} * 10 = 75$$

Quinto decil:

$$\frac{\mathbf{65*5}}{\mathbf{10}} = \mathbf{32,5}\ ; D_5 = 70 + \frac{32{,}5-18}{16} * 10 = 79{,}06$$

Sexto decil:

$$\frac{\mathbf{65*6}}{\mathbf{10}} = \mathbf{39}\ ; D_6 = 80 + \frac{39-34}{14} * 10 = 83{,}57$$

Séptimo decil:

$$\frac{\mathbf{65*7}}{\mathbf{10}} = \mathbf{45,5}\ ; D_7 = 80 + \frac{45{,}5-34}{14} * 10 = 88{,}21$$

Octavo decil:

$$\frac{\mathbf{65*8}}{\mathbf{10}} = \mathbf{52}\ ; D_8 = 90 + \frac{52-48}{10} * 10 = 94$$

Noveno decil:

$$\frac{\mathbf{65*2}}{\mathbf{10}} = \mathbf{13}\ ; D_9 = 90 + \frac{58{,}5-58}{5} * 10 = 101$$

3.3.3 Centil o percentil (P)

Se conoce como Centil o Percentil a la calificación que asigna por bajo el k por ciento de las puntuaciones de una distribución. Los Centiles son un caso poco común de Cuantiles.

Los percentiles son en total 99 valores que dividen el grupo de datos en 100 partes iguales, no necesariamente se tiene que elaborar todos los valores, pero si se lo puede realizar aleatoriamente. Los percentiles dan los valores proporcionados al 1%, al 2% y al 99% de los datos. El percentil P50 coincide con la mediana.

Si se quiere un resultado más exacto se puede utilizar la siguiente expresión (datos agrupados en frecuencias):

[3.11] $$\boldsymbol{P_k = L_i + \frac{1}{f_i}\left(\frac{kN}{100} - f_a\right)}$$

Dónde:

- k: Porcentaje de casos del centil.
- Li: Límite inferior de la puntuación donde se encuentra el centil.
- fi: Frecuencia de la puntuación donde se encuentra el centil.
- N: Tamaño del grupo
- Fa: Frecuencia acumulada hasta el límite inferior de la puntuación donde se encuentra el centil.

Ejemplo 3.13: Sea la siguiente distribución. Calcular el percentil 35 y 60.

Tabla 3.7

Cálculo de percentiles

Intervalos	**f_i**	**F_i**
[50, 60)	8	8
[60, 70)	10	18
[70, 80)	16	34
[80, 90)	14	48
[90, 100)	10	58
[100, 110)	5	63
[110, 120)	2	65
	65	

Percentil 35 $\boldsymbol{P35} = \frac{\boldsymbol{65*35}}{\boldsymbol{100}} = \boldsymbol{22,75}$

$$P_k = 70 + \frac{22{,}75 - 18}{16} * 10 = 72{,}97$$

Percentil 60 $\boldsymbol{P60} = \frac{\boldsymbol{65*60}}{\boldsymbol{100}} = \boldsymbol{39}$

$$P_k = 80 + \frac{39 - 34}{14} * 10 = 83{,}57$$

3.4 Medidas de dispersión

Las medidas de dispersión son ventajosas porque nos facilitan la información agregada que nos permite calificar la confiabilidad de nuestra medida de tendencia central. Como algunas veces los datos están muy dispersos, la posición central es menos representativa de los datos, como un todo, que cuando estos se agrupan más estrechamente alrededor de la media.

Ya que reflejan problemas especiales, las distribuciones son muy dispersas, por lo tanto debemos ser competentes de distinguir que representa esa dispersión antes de abordar los problemas. Estas medidas son utilizadas con frecuencia por profesionales, como por ejemplo los analistas financieros que están impacientados por la dispersión de los dividendos de una compañía que van comenzando con valores muy grandes a valores negativos. Esto revela un riesgo mayor para los asociados y para los acreedores.

Las medidas de dispersión más utilizadas son las siguientes: el rango, la varianza, la distribución estándar y el coeficiente de variación.

3.4.1 El rango (R)

Es la diferencia que hay entre el mayor y el menor de los valores observados.

[3.12] $$R = x_n - x_1$$

Siendo X_n la observación mayor y X_1 la observación menor.

El rango es fácil de entender y de encontrar, pero su utilidad como medida de dispersión es limitada. Como únicamente toma en cuenta el valor más alto y el valor más bajo, ignora la naturaleza de la variación entre las otras informaciones y se ve afectado por los valores de los extremos.

Debido a que se considera dos valores tiene probabilidad de cambiar rotundamente de una muestra a otra en una población dada. Además los grupos distribuciones de extremo abierto no tienen rango.

3.4.2 Varianza y desviación estándar

Las representaciones más perceptibles de la dispersión, son aquellas que tratan con la desviación promedio con respecto a alguna medida de tendencia central. Conoceremos dos medidas que nos dan una distancia promedio con respecto a la media de la distribución: la varianza y la desviación estándar.

Varianza de la población: Es el promedio de las distancias al cuadrado que van de las observaciones a la media.

Desviación estándar de la población: Es la raíz cuadrada de la varianza.

Cálculo de la varianza y la desviación estándar utilizando datos no agrupados

- σ 2: Varianza de la población
- x: Elemento u observación
- μ: Media de la población
- N: Número total de elementos de la población

[3.13] **Varianza:** $$\sigma^2 = \frac{\sum(x-\mu)^2}{N}$$

[3.14] **Desviación estándar:** $$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}}$$

Cálculo de la varianza y la desviación estándar utilizando datos agrupados

- σ 2: Varianza de la población
- σ: Desviación estándar de la población
- fi: Frecuencia absoluta de clase
- mi: Marca de clase
- μ: Media de la población
- N: Tamaño de la población

[3.15] **Varianza:** $\sigma^2 = \frac{\sum(m_i - \mu)^2 * f_i}{N}$

[3.16] **Desviación estándar:** $\sigma = \sqrt{\frac{\sum(m_i - \mu)^2 * f_i}{N}}$

Para calcular la **varianza y la desviación estándar muestral** se utilizan las mismas fórmulas que las poblacionales, sustituyendo **m** con **x** y N con **n-1**.

Ejemplo 3.14: La siguiente información constituyen una muestra de la cantidad de pedidos diarios entregados a una línea de los supermercados locales: 17; 25; 28; 27; 16; 21; 20; 22; 18; 23

Hallar lo siguiente: el rango, la varianza y la desviación estándar.

$\bar{x}$= 16; 17; 18; 20; 21; 22; 23; 25; 27; 28; Donde el promedio es 21,7.
$R = X_{10} - X_1$ = 28 - 16 = 12 La diferencia entre los valores observados del mayor y el menor da como resultado el
rango de 12

Para el cálculo de la varianza resulta conveniente realizar un cuadro como el expresado en la tabla 3.7, donde se detalla "X" como el rango de las observaciones que va desde 12 hasta 28 y $\bar{x}$ corresponde al promedio.

Tabla 3.8

Cálculo de la varianza

X	□	**X - □**	**(X-□)²**
16	21,7	-5,7	32,49
17	21,7	-4,7	22,09
18	21,7	-3,7	13,69
20	21,7	-1,7	2,89
21	21,7	-0,7	0,49
22	21,7	0,3	0,09
23	21,7	1,3	1,69
25	21,7	3,3	10,89
27	21,7	5,3	28,09
28	21,7	6,3	39,69
217		$\sum$	**152,1**

$$S^2 = \frac{\sum(X-\bar{X})^2}{N-1} = \frac{152{,}1}{10-1} = 16{,}9\ ;\ S = \sqrt{16{,}9} = 4{,}11$$

En promedio, la cantidad de pedidos se separa de la media, en 4,11 (pedidos).

3.4.3 Coeficiente de variación

El coeficiente de variación *CV,* es una medida relativa de dispersión que formula a la desviación estándar como una relación de la media. La desviación estándar es una medida absoluta de la dispersión, que formula la variación en las idénticas unidades que los datos originales. Pero no puede ser la única base para la comparación de dos distribuciones.

[3.17] $$CV_p = \frac{\sigma}{\mu} * 100\%\ en\ la\ población$$

[3.18] $$CV_m = \frac{s}{x} * 100\%\ en\ la\ muestra$$

Ejemplo 3.15: Supongamos una población de médicos, que trabajan ejerciendo su actividad en un país tropical, tienen un salario promedio de $968 y varianza de $12,000. Los galenos expertos que trabajan en un país de diferente trópico y en una actividad similar, tienen un salario promedio de $850 y una desviación típica (corresponde a la raíz cuadrada de la varianza) de $800. Se quiere determinar cuál grupo de salarios presenta una menor variabilidad.

Se calcula la varianza poblacional por medio de la fórmula.

$$CV_p = \frac{\sqrt{12{,}000}}{968} = 0{,}1132$$

$$CV_p = \frac{800}{850} = 0{,}94$$

Se obtiene como resultado que para la población del sector tropical posee un coeficiente de variación superior al de otros trópicos, lo que significa que el salario promedio en las regiones de clima tropical presenta menor variabilidad que en los otros trópicos.

Ejemplo 3.16: Si se analiza a un componente de toda la población una muestra de 4000 doctores de clima tropical cuyo ingreso promedio asciende a 950 y una desviación típica de 125, se obtiene:

$$CV_m = \frac{125}{950} = 0{,}1315$$

Al realizar el análisis muestral se encuentra que posee una mayor variabilidad 13,15% respecto a 11,32% de la población. Ello se explica a que se determina el análisis de un componente de toda la población.

3.4.4 Interpretación de las medidas de dispersión

Las medidas de dispersión tienen las siguientes funciones:

a) Calificar al promedio, es decir indicar si el promedio es representativo de la serie. Que sea representativo de la serie significa que el "promedio", es un indicador de la mayor parte de los datos, es decir que a través de ese número puedo sacar conclusiones de todos ellos.

Si una medida de dispersión es muy alta, o en extremos casos, mayor al promedio, ello indica que el promedio utilizado no es representativo de los datos observados y por lo tanto debe descartarse ese indicador y en su lugar elegir otro promedio. Dado que el promedio utilizado en el cálculo de la dispersión es la Media Aritmética, el indicador que debe ser reemplazado por otro en caso de tener una alta dispersión, es la Media Aritmética por la Mediana, dado que ésta última, no se ve afectada por valores muy altos o bajos.

b) Especificar el grado de alejamiento o acercamiento de los datos con respecto a su promedio. Esto permite determinar los valores máximos y mínimos de la serie, es decir el Rango, que permitiría tomar decisiones al Administrador con respecto a temas de calidad, de aceptar o no una determinada obra que tiene especificaciones técnicas, de conceder o no un aumento de sueldo, de determinar si un grupo de estudiantes cumple o no con un requerimiento específico para otorgamiento de una beca.

Las medidas de dispersión generalmente se las utiliza para medir el comportamiento de requisitos o especificaciones previamente definidas, donde el parámetro "promedio", sólo es uno de los elementos para la toma de decisiones, y más bien la "dispersión", es aquel que define su comportamiento real.

c) Las medidas de dispersión indican la variabilidad de los datos, es decir, en qué grado los datos no son estables, referidos a un promedio, de allí que son de suma importancia, en el caso de que se requiera tomar decisiones respecto a tasas de rentabilidad de inversiones que se requiere sean estables, si se desea que el "promedio de interés "obtenido no tenga altibajos, una alta dispersión de una tasa de rentabilidad se asemeja a un riesgo alto tanto de perder como de ganar altas sumas de dinero.

3.5 Medidas de forma

En las medidas descriptivas de variabilidad o también conocidas como de forma, existen dos características de los grupos de datos que proporcionan información útil: el sesgo (asimetría) y la curtosis.

3.5.1 Sesgo (asimetría)

Las curvas que constituyen el grupo de datos pueden ser simétricas o sesgadas. Las curvas simétricas poseen una forma tal que una línea vertical que pase por el punto más alto de la curva, fracciona al área de ésta en dos partes semejantes. Si los valores se agrupan en un extremo se dice **sesgada,** tanto si esta tiene sesgo positivo que es cuando los valores van disminuyendo lentamente hacia el extremo derecho de la escala y sesgo negativo en caso contrario.

- El sesgo es una medida de la asimetría de la curva.
- En general es un valor que va de -3 a 3.
- Una curva simétrica toma el valor 0.

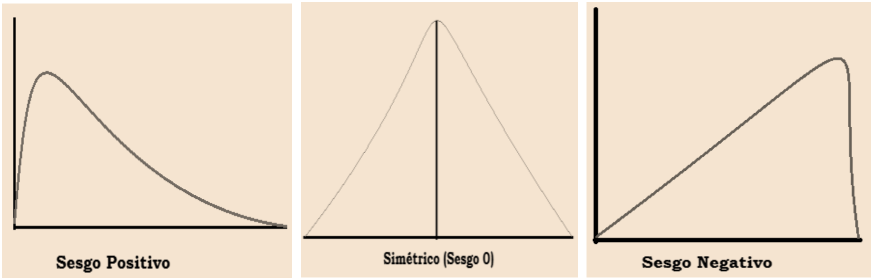


Ilustración 3.2 Tipos de sesgos

3.5.2 Curtosis (K)

Esta medida proporciona una idea de lo sobresaliente (o lo plano) de la distribución de frecuencias. Es una curva normal (es el modelo con el que se asemeja la curtosis de otras curvas) tiene curtosis 0 esta curva se llama mesocúrtica. Si la curtosis es mayor que 0, la curva es más levantada que la primera y se denomina leptocúrtica (Lepto, del griego, "empinado" o "estrecho"). Si la curtosis es menor que 0, es respectivamente plana y se denomina platicúrtica ("plano", "ancho").

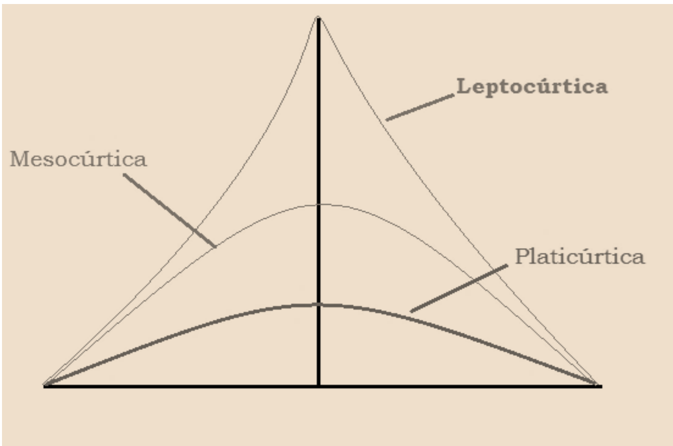


Ilustración 3.3 Tipos de curtosis

Por lo tanto se define como curtosis la medida que analiza el grado de concentración de los valores de una variable alrededor del límite central de las distribuciones de frecuencia.

Cálculo de la curtosis

[3.19] $$K = \frac{\sum_{i=1}^{n} n(x_i - \bar{x})^4}{nS^4} - 3$$

Donde:

x_i= cada uno de los valores;

n = número de datos;

$\bar{x}$= media aritmética;

S^4= Cuádruplo de la desviación estándar poblacional.

3.6 Análisis exploratorio de datos (A.E.D.): Unidimensional

El análisis exploratorio de datos (A.E.D.) es un agregado de técnicas estadísticas cuya finalidad es lograr una comprensión básica de los datos y de las relaciones entre las variables analizadas. Para obtener este objetivo, el A.E.D. facilita métodos consecuentemente sencillos para organizar y acomodar los datos, mostrar fallos en el diseño y recolección de los mismos, procedimiento y evaluación de datos faltes, reconocimiento de casos atípicos y confirmación de los supuestos subyacentes en la gran parte de las técnicas multivariantes (normalidad, linealidad, homo elasticidad).

3.6.1 Etapas del análisis exploratorio de datos

Para formalizar un análisis exploratorio de datos, es conveniente seguir las siguientes etapas:

- Acomodar los datos para brindar accesibles a cualquier técnica estadística.
- Perpetrar un examen gráfico de la naturaleza de las variables individuales a examinar y un estudio descriptivo

numérico que admita cuantificar algunos aspectos gráficos de los datos.

- Efectuar un exámen gráfico de las interacciones entre las variables analizadas y un análisis descriptivo numérico donde se cuantifique el grado de interrelación existente entre ellas.
- Identificar los potenciales casos atípicos y valorar el impacto potencial que puedan ejecutar en los análisis estadísticos posteriores.
- Evaluar, el impacto potencial que pueden tener los datos faltes sobre la representatividad de los datos analizados.

3.6.2 Análisis estadístico unidimensional

Una vez establecidos los datos, el segundo paso de un A.E.D. consiste en efectuar un estudio estadístico gráfico y numérico de las variables del problema con el resultado de tener una idea inicial de la información agrupada en el conjunto de datos así como descubrir la existencia de potenciales errores en la categorización de los mismos.

La clase de análisis a realizar obedece de la escala de medida de la variable analizada. A veces se proponen las representaciones gráficas y resúmenes descriptivos numéricos más recomendables para realizar dicho análisis.

3.6.3 Variables cualitativas

Los datos proporcionados a variables cualitativas se agrupan de manera natural en diferentes categorías o tipología y se calcula el número de datos que surgen en cada una de ellas. Se suelen mostrar mediante diagrama de barras, sectores o líneas.

Ejemplo 3.17: En una encuesta realizada en un supermercado en la siguiente tabla se muestra la tabla de frecuencias del Estado Civil de una muestra extraída de los clientes del Supermercado. Así mismo, en la Figura se muestra el diagrama de sectores correspondiente a esta variable.

Tabla 3.9

Frecuencias de la encuesta

Tabla de Frecuencia del Estado Civil		
Variable	Frecuencia	Porcentaje
Soltero	77	19,2
Casado	305	75,9
Viudo	16	4
Divorciado	4	1
Total	**402**	**100**

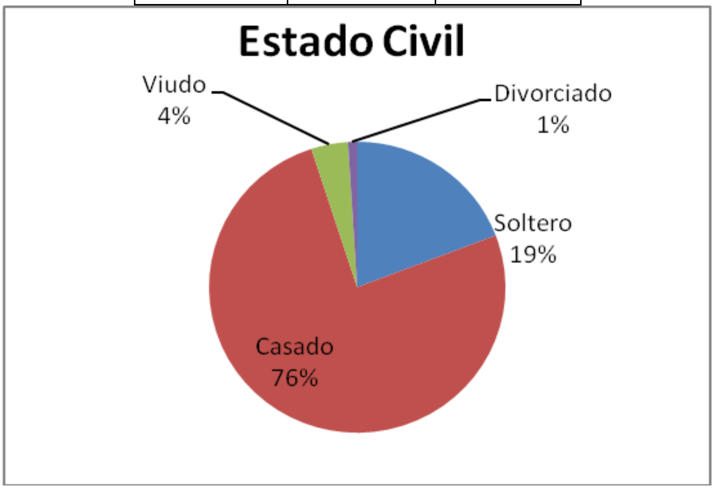


Ilustración 3.4 Gráfica del estado civil de las personas

3.6.4 Variables cuantitativas

Las variables cuantitativas discretas con un número pequeño de valores se tratarían de manera similar a las variables cualitativas antes descritas.

Ejemplo 3.18: En la siguiente tabla se muestra la distribución de frecuencias del número de miembros que viven en la casa de una muestra de clientes de un supermercado. Así mismo, la tabla muestra ciertas medidas descriptivas numéricas de dicha distribución y su diagrama de barras.

Tabla 3.10

Tabla indicadora de frecuencias

		fi	Fi	Porcentaje Válido	Porcentaje Acumulado
Válidos	0	1	1	0,3	0,3
	1	30	31	7,5	7,8
	2	91	122	22,8	30,5
	3	87	208	21,8	52,3
	4	129	337	32,3	84,5
	5	43	380	10,8	95,3
	6	12	392	3	98,3
	7	7	400	1,8	100
	Total	400		100	
Perdidos	Sistema	2			
Total		402			

Tabla 3.11

Indicadores paramétricos

Miembros de un hogar		
N	Validos	400
	Perdidos	2
Media		3,31
Mediana		3
Moda		4
Desv. Típ.		1,33
Asimetría		0,234
Error típ.asim.		0,122
Curtosis		-0,107
Error típ.curt.		0,243
Mínimo		0
Máximo		7
Percentiles	25	2
	50	3
	75	4

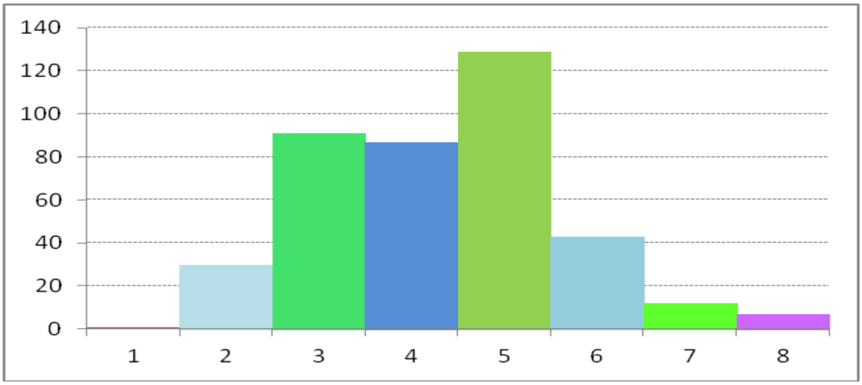


Ilustración 3.5: Gráfico de frecuencias absolutas

3.7 Ejercicios propuestos

1.- La Corporación Financiera Nacional (CFN) contrata a veinte profesionales que iniciaron en el trabajo con los siguientes sueldos ($): 843,50; 843,50; 908,00; 856,00; 695,00; 660,00; 605,00; 903,50; 980,00; 869,00; 695,00; 750,00; 660,00; 980,00; 680,00; 843,50; 576,00; 925,00; 919,00; 843,50.

- Calcular la media, mediana y la moda.
- ¿Cuál de los promedios anteriores es el más representativo y por qué?

2.- Un grupo de 100 obreros en una mina, 60 trabajan en el día y 40 en la noche. Se sabe que el salario promedio de los cien obreros es de $350,00 y que los del turno de día reciben en promedio.
$58,00 menos que los trabajadores nocturnos. ¿Cuál es el salario promedio en cada grupo?

3.- Calcular la media de la siguiente tabla distribución estadística:

Intervalos	**fi**
[0, 5)	3
[5, 10)	5
[10, 15)	7
[15, 20)	8
[20, 25)	2
[25, ∞)	6

4.- Calcula la media aritmética ponderada encontrando la nota promedio de dos materias que tienen calificaciones diferentes. Estadística tiene un valor de 3 créditos y álgebra vale 2 créditos. Si Carlos ha sacado un 8 en la primera y un 7 en la segunda.

5.- Se calcula la nota final del curso de literatura en donde cada nota ha tenido distinta importancia. Los dos primeros trabajos tienen valor de 20% y 20% respectivamente y el examen de 60%; las calificaciones respectivas son de 6.4, 9.2 y 8.1. Hallar la media ponderada del curso.

6.- En un almacén las devoluciones mensuales son a lo mucho el 10 % tienen un importe de $ 360. El 40 % son de un importe inferior o igual a $660. En la mitad de ellas no se superan las 1385 mientras que el 30 % están entre $ 1385 y $ 2410

- Si se sabe que en dicho almacén el importe máximo de las devoluciones es de $3000, calcula e interpreta el importe medio de las devoluciones.
- ¿Podrías calcular el importe medio si no se supiera cuál es el importe máximo de las devoluciones?

7.- En un intercolegial, un grupo de estudiantes se considera el número de ensayos que necesita cada uno para aprender una lista de seis pares de palabras. Los resultados fueron:
5, 8, 3, 9, 6, 7, 10, 6, 7, 4, 6, 9, 5, 6, 7, 9, 4, 6, 8, 7, 9, 6, 7, 10, 8, 1, 3, 5, 7 , 4, 2, 7, 8, 9, 10, 2, 4, 3, 4;

a) Construya la tabla de frecuencias.

b) Hallar la moda, la media, la mediana, el segundo y tercer cuartil de las observaciones dadas. Obtenga la frecuencia del conjunto de los resultados superiores a 5.

c) Hallar la varianza y el desvío estándar.

8.- Un estudio que se efectuó en un hospicio de ancianos, se tomó las edades de los adultos mayores que pueden caminar sin dificultades. Buscar la media, la mediana y la moda de las siguientes edades.

69 73 65 70 71 74 65 69 60 62 65 80 94 68 83 75 65 69 65 65 69 78 89 80 64 89 65 74 77 73 70 78 62 65 94 69 75 89 89 64 73 65 89.

9.- Dada la siguiente distribución:

Xi 2 3 8 12 17 9 6 11 10 7 15 4 1

Ni 2 2 3 3 1 3 5 3 3 2 6 4 4

Calcular:

- La media aritmética.
- La media geométrica.
- La media armónica.
- Comprobar la relación que existe entre ellas.

10.- Las calificaciones de la materia de Estadística del curso A3 se distribuyen de acuerdo a la siguiente tabla para los alumnos presentados en junio:

Calificación	**Valor**	**Alumnos**
Mala	0	2
Regular	1	10
Buena	2	90
Muy Buena	3	123
Sobresaliente	4	12

- Represente la gráfica de frecuencias.
- Calcule la media aritmética, la moda y la mediana.
- Calcule la varianza y el cuartil Q 3/4.

3.8 Ecuaciones introducidas en el capítulo 3

Media aritmética muestral
Una muestra con n observaciones, tiene una media χ denominada media estadística, del elemento de una muestra.

[3.1] $$\boldsymbol{x = \frac{\sum x}{n}}$$

Media aritmética poblacional
Una población con N elementos tiene una media μ denominada parámetro de los datos de una población completa

[3.2] $$\mu = \frac{\sum \boldsymbol{X}}{\boldsymbol{N}}$$

Media aritmética ponderada (datos agrupados)
Cuando el número de observaciones es grande, las operaciones para calcular la media aritmética se simplifican si agrupamos los datos en una tabla de frecuencias.

[3.3] $$\boldsymbol{x = \frac{\sum m_i f_i}{n}}$$

Mediana con datos no agrupados
Número impar de datos: La mediana es el dato que está en la posición $\boldsymbol{\frac{n+1}{2}}$

[3.4] $$\boldsymbol{Me = X_{(\frac{n+1}{2})}}$$

Número par de datos: Es el promedio entre los dos datos centrales.

[3.5] $$\boldsymbol{Me = \frac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2}}$$

Mediana con datos agrupados
Si los datos están concentrados en una distribución de frecuencias, se elige el intervalo de clase que domina a la mediana conocido como clase mediana.

[3.6] $$Me = X = L_m + \left[\frac{\frac{n+1}{2} - F_{i-1}}{f_i} \right] * h$$

Media geométrica con datos no agrupados

El promedio geométrico se calcula hallando el producto de todos los elementos de la serie, y luego extrayendo la raíz del orden del número de observaciones consideradas.

[3.7] $$Mg = \sqrt[n]{\prod_{i=1}^{n} x_i}$$

Media geométrica datos agrupados

La media geométrica se define como la raíz enésima del producto de los valores de la variable, elevadas cada una de ellas a una potencia, la cual está dada por la frecuencia absoluta.

[3.8] $$Mg = \sqrt[n]{\prod_{i=1}^{n} y_i^n}$$

Reglas logarítmicas para el análisis de media geométrica con datos agrupados

[3.9] $$LogMg = \frac{\sum n_i logy}{n}; Mg = anti\ log \frac{\sum n_i logy}{n}$$

Media armónica

El promedio armónico se precisa expresando que el recíproco de la media armónica es igual a la media aritmética del recíproco de los valores de la variable.

[3.10] $$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{Xi}}$$

Centil o percentil

Se conoce como Centil o percentil a la calificación que asigna por bajo el k por ciento de las puntuaciones de una distribución. Los Centiles son un caso poco común de Cuantiles.

[3.11] $$P_k = L_i + \frac{1}{f_i}(\frac{kN}{100} - f_a)$$

El rango
Es la diferencia que hay entre el mayor y el menor de los valores observados.

[3.12] $$\boldsymbol{R = x_n - x_1}$$

Cálculo de la varianza y la desviación estándar utilizando datos no agrupados

[3.13] **Varianza:** $$\boldsymbol{\sigma^2 = \frac{\sum(x - \mu)^2}{N}}$$

[3.14] **Desviación estándar:** $$\boldsymbol{\sigma = \sqrt{\frac{\sum(x - \mu)^2}{N}}}$$

Cálculo de la varianza y la desviación estándar utilizando datos agrupados.

[3.15] **Varianza:** $$\boldsymbol{\sigma^2 = \frac{\sum(m_i - \mu)^2 * f_i}{N}}$$

[3.16] **Desviación estándar:** $$\boldsymbol{\sigma = \sqrt{\frac{\sum(m_i - \mu)^2 * f_i}{N}}}$$

Coeficiente de variación

[3.17] $$CV_p = \frac{\sigma}{\mu} * 100\% \; en\ la\ población$$

[3.18] $$CV_m = \frac{s}{x} * 100\% \; en\ la\ muestra$$

Curtosis (K)
Esta medida proporciona una idea de lo sobresaliente (o lo plano) de la distribución de frecuencias.

[3.19] $$\mathrm{K} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{n}(x_i - \bar{x})^4}{\mathrm{n}S^4} - 3$$

Unidad 4

NÚMEROS ÍNDICES

"El desarrollo es más que un número".

-AMARTYA SEN-

Premio Nobel de Economía.

CONTENIDO DE LA UNIDAD

INTRODUCCIÓN

En este capítulo se estudiará cómo se mide el cambio de precios o cantidades de un grupo de artículos que intervienen en la economía, por lo tanto se explicará los métodos que se usan para elaborar los números índices, ya sea de manera simple, debido a que abarca el análisis de un solo concepto, como sería el caso de consultar el índice de precios al consumidor IPC citados en las noticias de un medio de comunicación local, o de forma compleja, por que abarca varios conceptos, como sería el caso del cálculo del IPC que analiza las variaciones de las producciones asociadas a ella, con ponderaciones y sin ella mostrando los indicadores necesarios para la interpretación de estos resultados.

OBJETIVOS DE APRENDIZAJE

Al finalizar el capítulo, será capaz de:

- Identificar los tipos de índices más utilizados, ponderados y no ponderados.
- Comprender la diferencia entre un índice ponderado y uno no ponderado.
- Reconocer los elementos que intervienen cuando se elabora el Índice.
- Analizar e interpretar los resultados de los índices, valor y precio.
- Determinar los problemas en la construcción y uso de los números índices.

Es importante resaltar que los números índices son utilizados en las ciencias económicas, ya que son una manera para generar modelos de comportamiento económico, relacionados con precios, cantidades, producción etc.

4.1 Definición de números índices

4.1.1 Números índices

Definición: Un índice o número índice calcula el cambio en un concepto en particular (un producto o servicio) entre dos períodos, es el número que expresa el cambio referente en precio, cantidad o valor comparado con un período base.

Un índice representa el cambio relativo de un valor perteneciente de un período a otro. Sin duda, se conocen índices como el Índice de Precios al Consumidor pero también hay muchos índices, como el Industrial Average (DJIA), promedio Dow Jones, el Nasdaq, Standard & Poor.

El gobierno estadounidense informa índices de manera reiterada en revistas de negocios como Business Week y Forbes, en la mayoría de los medios de comunicación y acceso de internet.

Elaboración de número índice

Así se elabora un índice simple: el valor en un año seleccionado, se divide entre el precio del año base. El valor en el período base se designa ***I(n)*** y un precio que no sea el período base se conoce como período dado o seleccionado y se designa ***I(n+1)***. Para calcular este índice de precios simple P con 100 como valor base para un período dado, utilizar la fórmula:

[4.1] $$I = \frac{I(n+1)}{I(n)} * 100$$

Ejemplo 4.1: Conforme con el INEC, en 1990 el salario promedio por hora de los médicos era \$9,3; en 1995 fue de \$6,5; en el 2000 de \$9,6 y el 2005 de \$10,1. ¿Cuál es el índice de salarios por hora de los médicos para el 2005 con base de 1990?

Tabla 4.1

Índice de salarios

Año	Salario hora	Cociente	Índice o Porcentaje
1990	9,3	1	100
1995	6,5	0,7	70
2000	9,6	1,03	103
2005	10,1	1,09	109

$$I = \frac{Salario\ por\ hora\ 2005}{Salario\ por\ hora\ 1990} * 100$$

$$I = \frac{10{,}1}{9{,}3} * 100 = 109$$

Efectivamente, el salario por hora en el 2005 comparado con el año 1995 fue 109%. Esto representa que hubo un crecimiento de 9% en el salario por hora durante el período, (109 – 100.0 = 9%).

4.2 Tipos de números índices

Además podemos recurrir a los números índices para detallar cambios en cantidades y en valores:

4.2.1 Índices de precios (P)

Los índices de precios estudian la evolución de los precios de un bien o de un conjunto de bienes. El IPC es un indicador cuyo objetivo es calcular la variación y cambio de nivel de los precios del grupo de bienes y servicios adquiridos o consumidos por los hogares (personas) en una economía. El índice más utilizado es el índice de precio de consumo (IPC), este permite identificar las variaciones de los precios con un período base.

4.2.2 Índice de cantidad (Q)

Un índice de cantidad facilita una medida más confidencial de la producción real de materias primas y bienes terminados que el conveniente índice de valores. De modo parecido, la producción agrícola se calcula mejor si se realiza un índice de cantidad, debido a que éste descarta los efectos exagerados producidos por la variación de precios.

4.2.3 Índices de valores (V)

Los índices de valores estudian el progreso del valor de un bien o de un grupo de bienes. El valor de un grupo de bienes y/o servicios, para dos ciclos de tiempo, la vigente t y la base 0, vendrá dado proporcionalmente por las siguientes expresiones:

$\boldsymbol{V_T} = \sum \boldsymbol{V_{it}} = \sum(\boldsymbol{p_{it}} * \boldsymbol{q_{it}})$ (Valor en el período actual)

$\boldsymbol{V_0} = \sum \boldsymbol{V_{i0}} = \sum(\boldsymbol{p_{i0}} * \boldsymbol{q_{i0}})$ (Valor en el período base)

4.3 Beneficios del uso de números índices

Los números índices pueden utilizarse de varias maneras. Lo más común es usarlos por sí mismos, como un resultado final. Los números índices, como el IPC, a menudo se citan en informes noticiosos como indicadores generales de la condición económica de un país.

La gerencia o administración, utiliza los números índices para formar un cálculo intermedio permitiendo comprender mejor otra información. El uso del IPC determina el poder adquisitivo real del dinero, esto es otro ejemplo de cómo los índice ayudan a aumentar el conocimiento de otros factores.

4.4 Limitaciones a considerar en el manejo de números índices

Existen cuatro factores que pueden alterar los números índices, estas cuatro causas que afectan los números índices son:

1. **La ponderación a usar, no es apropiada de los factores,** se puede alterar un índice. Al momento de desarrollar un número índice compuesto, por ejemplo el IPC, es preciso tomar en cuenta que los cambios en cualesquiera de las variables son más significativos que en otros.
2. **Dificultad en encontrar datos adecuados para desarrollo de un índice**. En el supuesto que el administrador de una empresa desea calcular un índice que represente la variación de las ventas de la compañía y estas se registran sólo anualmente, el administrador no será capaz de determinar el patrón de ventas estacional.
3. **La falta de igualación de índices,** esto sucede cuando se compara un índice con otro después de que ha tenido una variación en lo que se mide.
4. **La distorsión de los números índices** también ocurre cuando **se selecciona una base no apropiada.** A veces, una compañía selecciona una base que automáticamente conduce a un resultado que refleja sus propios intereses y lo usa para probar su suposición inicial.

4.5 Números índices no ponderados

Índices sin ponderar: Este número índice es la forma sencilla perteneciente al índice de agregados no ponderados. Cuando los datos son no ponderados representa que todos los valores supuestos tienen la misma importancia; agregados

se refiere a la sumatoria de todos los valores. La mayor ventaja de un índice de agregados no ponderados es su simplicidad.
El índice de agregados no ponderados se realiza la sumatoria todos los elementos del compuesto para el período a analizar, para luego dividir este resultado entre la sumatoria de los elementos en el período base. La ecuación 4.2 demuestra la fórmula matemática para desarrollar un índice de cantidad de agregados no ponderados.

[4.2] $$\boldsymbol{I(x)} = \frac{\sum \boldsymbol{Q_I}}{\sum \boldsymbol{Q_0}} * 100$$

Donde:

Q_i: número de variables (precio, cantidad o valor) o elementos en el año para el que se desea el índice.

Q_0: número de variables de cada elemento del compuesto en el año base.

Ejemplo 4.2: El INEC desea saber la variación de precios del los productos básicos del desayuno de la mesa ecuatorian. En este caso, queremos medir los cambios en los niveles generales de precios con base en los cambios de precios de unos cuantos artículos. Los precios para 2007 son los valores base con los que se comparan los precios del 2017.

Tabla 4.2

Datos de la variación de precios

Productos Básicos	**Precios**	
	2007	**2017**
Leche (litro)	0,75	1,25
Huevos (docena)	1	1,5
Pan (unidad)	0,1	0,22
Azúcar (kl.)	0,5	1,25
	2,35	**4,22**

$$I(x) = \frac{4{,}22}{2{,}35} * 100 = 179{,}5\%$$

Esto indica que el precio de venta medio del 2017, es 179,5% del precio de venta medio del 2007. En otras palabras, el precio de venta medio es 79,5%

(179,5 – 100,0 = 79,5%).

4.6 Números índices ponderados

Índices ponderados: Cada dimensión o componente tiene un valor diferente asignado en función de diversos criterios. Los k datos se consideran con distinto peso, peso que acumula la importancia relativa de cada uno de los bienes. Esta ponderación permite contener más información, que sólo la alteración de los precios en el tiempo, nos permite mejorar la precisión de la estimación del nivel general de precios basado en una muestra.

[4.3] $$\boldsymbol{I_P = \frac{\sum P_I Q}{\sum P_0 Q} * 100}$$

Donde:

P_i: precio de cada elemento del compuesto en el año actual.

P_0: precio de cada elemento del compuesto en el año base.

Q: factor de ponderación de cantidad seleccionado.

4.6.1 Método de Laspeyres

Conocida como la media aritmética ponderada de índices simples de cada concepto usándose como ponderación para cada bien $\mathbf{I_i = p_i.q_0}$, esto es la ponderación para cada concepto será el valor de la cuantía consumida, vendida o producida del bien i-simo en el momento base al precio del período base.

La desventaja principal del método de Laspeyres es que no toma en cuenta los cambios en los patrones de consumo.

[4.4] $$\boldsymbol{L_P = \frac{\sum P_I Q_0}{\sum P_0 Q_0}} * 100$$

Donde:

P_i: precio de cada elemento del compuesto en el año actual.

P_0: precio de cada elemento del compuesto en el año base.

Q_0: cantidades vendidas en el año base.

Ejemplo 4.3: Hallar el Índice de Laspeyres para las cantidades y precios

Tabla 4.3
Índice de Laspeyres

Productos Básicos	Cantidad Año 2007 (Q_0)	Precios (P)		Vtas. Ponderadas (P_0Q)	Vtas. Ponderadas (P_iQ)
		2007 (P_0)	2017 (P_i)		
Leche (litro)	300	0,75	1,25	225	375
Huevos (docena)	50	1	1,5	50	75
Pan (unidad)	600	0,1	0,22	60	132
Azúcar (kl.)	140	0,5	1,25	70	175
		2,35	**4,22**	405	757

$$L_P = \frac{757}{405} * 100 = 186{,}92\%$$

Esto indica que la cantidad medio del 2017, es 186,92% de la cantidad medio del 2007. En otras palabras, el precio de venta medio es 86,92%

(186,92 – 100,0 = 86,92%).

4.6.2 Método de Paasche

Representa la media aritmética ponderada de los números índices simples de cada concepto usándose como ponderación para cada artículo: wi = pi0.qit, esto es, el valor a precio del momento base de la cantidad consumida en el período determinado como el actual.

[4.5] $$P_P = \frac{\sum P_i Q_i}{\sum P_0 Q_i} * 100$$

Donde:

Pi: precio de cada elemento del compuesto en el año actual.

P0: precio de cada elemento del compuesto en el año base.

Qi: cantidades vendidas en el período actual.

Ejemplo 4.4: Con datos del ejemplo anterior (Ejemplo 4.3) hallar el índice de Paasche.

Tabla 4.4

Índice de Paasche

Productos Básicos	Precios (P)		Año 2017 (Q_i)	Vtas. Ponderadas (P_iQ_i)	Vtas. Ponderadas (P_0*Q_i)
	2017 (P_i)	**2007 (P_0)**			
Leche (litro)	1,25	0,75	500	625	375
Huevos (docena)	1,5	1	75	112,5	75
Pan (unidad)	0,22	0,1	1000	220	100
Azúcar (kl.)	1,25	0,5	100	125	50
	4,22	**2,35**		1082,5	600

$$L_P = \frac{1082{,}5}{600} * 100 = 55{,}43\%$$

En este análisis, encontramos que el índice de precios para el 2017 es 55,43%. Como se ve en el desarrollo del ejercicio 4.4, el índice de precios analizados mediante el método de Laspeyres es 186,92%. La diferencia entre estos índices muestra el cambio en los estándares de consumo.

El método de Paasche es especialmente ventajoso porque adopta los efectos de las variaciones de precio y los modelos de consumo. Así, es un mejor indicador de las variaciones generales de la economía.

4.6.3 Método de agregados con peso fijo

Este método posee una técnica para determinar pesos a los elementos de de estudio. Es parecido a los métodos anteriormente vistos, de Laspeyres y Paasche, pero en cambio de manejar pesos del período base o un período en curso (cantidades), usa pesos seleccionados de un período representativo. Los pesos representativos se conocen como pesos fijos. Estos pesos fijos y los precios base no tienen que corresponder al mismo período.

[4.6] $$\boldsymbol{F_P = \frac{\sum P_iQ_2}{\sum P_0Q_2} * 100}$$

Donde:

P_i: precio de cada elemento en el año actual.

P_0: precio de cada elemento en el año base.

Q_2: pesos fijos.

Ejemplo 4.5: La empresa Multi-Metal Steel fabrica partes de calidad para la producción de máquina pesada. Los principales materiales que utiliza son carbón, minerales de hierro, níquel y cobre. La gerencia tiene los siguientes datos de los costos de estos materiales en 2000 y 2017, así como la cantidad de materiales utilizados en 2010, cuando patrones de adquisición fueron propios de todo el período de 17 años.

Tabla 4.5

Producción de materiales

Materia Prima (ton.)	Consumo 2010 (Q_2)	Precios (P)		Agregado Ponderado (P_0Q_2)	Agregado Ponderado (P_i*Q_2)
		2000 (P_0)	2017 (P_i)		
Carbón	180	$ 8,25	$ 10,60	1.485,00	1.908,00
Hierro	12	$ 12,30	$ 13,75	147,60	165,00
Níquel	40	$ 13,75	$ 15,50	550,00	620,00
Cobre	6	$ 20,45	$ 25,25	122,70	151,50
				2305,3	2.844,50

$$F_P = \frac{2844,5}{2305,3} * 100 = 123,39\%$$

Posteriormente de reconocer los registros de compras, la gerencia decidió que las cantidades adquiridas en 2000 fueron características de los modelos de adquisición durante los 17 años. El análisis del índice de agregados con peso fijo se origina en la tabla 24. La gerencia de la compañía detalla, que a partir de este análisis, los precios ha aumentado el 123,39% en el período de 17 años.

4.7 Métodos de promedios de relativos

4.7.1 Método de promedio no ponderado de relativos

Cuando tenemos más de un objeto (o actividad), primero hallamos el cociente del precio vigente entre el precio base para cada objeto y multiplicamos cada cociente conseguido por 100. Luego sumamos los porcentajes relativos resultantes y dividimos el resultado con el número de productos.

$$x_R = \frac{\sum(\frac{P_i}{P_0} * 100)}{n} \qquad [4.7]$$

Donde:

Pi: precios del período actual**.**

P0: precios del período base.

n: número de elementos o productos del compuesto.

Con este método deducimos el promedio de los cocientes de los valores de los precios para cada objeto. Con el método de agregados no pesados, calculamos el cociente de las sumas de los precios de cada producto.

No es igual asignar a algunos objetos más peso que a otros. El método de promedio de relativos cambia cada elemento a una sucesión relativa en la que los elementos están personificados como un porcentaje más que como una cantidad. Obligado a esto, cada uno de los elementos del compuesto se calcula con relación a una base de 100.

Ejemplo 4.6: Con datos del ejemplo anterior (Ejemplo 4.5) realizar el método de promedio no ponderado relativo.

Tabla 4.6

Cálculo de método de promedio no ponderado

Materia Prima (ton.)	Precios (P)		Cociente Pi/P0*100
	2000 (P0)	2017 (Pi)	
Carbón	$ 8,25	$ 10,60	128,48
Hierro	$ 12,30	$ 13,75	111,79
Niquel	$ 13,75	$ 15,50	112,73
Cobre	$ 20,45	$ 25,25	123,47
			476,47

Al momento de utilizar la formula tenemos:

$$x_R = \frac{476{,}47}{4} = 119{,}12$$

El método de promedio no ponderado de relativos se lo calcula mediante un promedio de los cocientes de los precios para cada producto. Se debe considerar que esto no es lo mismo que fijar más peso a cualesquiera productos que a otros. Más apropiadamente, el método de promedio de relativos cambia cada componente a una escala relativa en donde los elementos se personifican como un proporción y no como una cantidad.

4.7.2 Método de promedio ponderado relativos

Los métodos de promedio ponderado de relativos permiten de varias maneras, establecer un valor con peso relativo. Como en el método de Laspeyres, suele manejar el valor base que hayamos multiplicando la cantidad base por el precio base. La utilización del valor base provocará exactamente el mismo resultado que si estuviéramos deduciendo el índice con el método de Laspeyres. Empleamos el método Laspeyres cuando la información de la cantidad se obtiene con mayor facilidad.

[4.8] $$\boldsymbol{x_R = \frac{\sum(\frac{P_i}{P_0} * 100)(P_n Q_n)}{\sum P_n Q_n}}$$

Donde:

Pi: precios del período actual**.**

P0: precios del período base.

Pn: precios que determinan los valores que utilizamos para los pesos.

Qn: cantidades que determinan los valores que utilizamos para los pesos.

Cuando manejamos valores actuales, no podemos medir de manera directa valores numéricos de diferentes períodos, sea por que tanto los precios como las cantidades logran haber cambiado. Por lo general recurrimos a valores base o valores fijos cuando calculemos un índice de promedio pesado de relativos.

Ejemplo 4.7: Con datos del ejemplo anterior (Ejemplo 4.5) realizar el método de promedio ponderado relativo.

Tabla 4.7

Método de promedio ponderado

Materia Prima (ton.)	Precios (P)		Cantidad 2000 (Q0)	Porcentajes de Precios Relativos	Valor Base (P0*Q0)	Pondrado Relativo
	2000 (P0)	2017 (Pi)				
Carbón	8,25	10,6	200	128,48	$ 1.650,00	$ 212.000,00
Hierro	12,3	13,75	300	111,79	$ 3.690,00	$ 412.500,00
Níquel	13,75	15,5	100	112,73	$ 1.375,00	$ 155.000,00
Cobre	20,45	25,25	150	123,47	$ 3.067,50	$ 378.750,00
					$ 9.782,50	$1.158.250,00

Al momento de desarrollar la fórmula tenemos:

$$x_P = \frac{1.158.250}{9.782,5} = 118,40$$

La información de la tabla 4.6 viene de la tabla 4.5. Tenemos precios base y cantidades base, de modo que usaremos la ecuación [4.8]. El índice de precios 118,4.

4.8 Ejercicios propuestos

1.- Un 30 de enero del 2015 empresa liquidó una nómina de 3 millones de dólares a sus 88 trabajadores. El 30 de Enero del siguiente año, dicha empresa aumentó en 10 sus trabajadores y liquidó una nómina de 4,3 millones de dólares más que el año anterior. Halle:

(a) El índice del número de trabajadores para Enero de 2016.

(b) El índice del gasto en nómina para Enero de 2016.

(c) El índice del costo medio por trabajador para Enero de 2016.

2.- En año 2010 el precio del quintal de arroz bajó un 20% con proporción al año anterior, pero aumentó un 50% con relación al año 2005. Halle el número índice del precio de 2009 tomando como año base 2005.

3.- En una compañía los costes de producción son:

- Sueldo de obreros 45%
- Materias primas 30%
- Gastos de administración 10%

Los sueldos se adecuan al índice de coste de la vida que ha pasado de 150 a 168 en los posteriores 5 años, en este lapso de tiempo las materias primas se han elevado un 8% y los gastos de administración no han variado. Hallar el incremento del coste de producción en este periodo de tiempo.

4.- Conocidos los precios y cantidades de tres artículos de consumo correspondientes a los años 2001 a 2005:

ARTÍCULOS	2001		2002		2003		2004		2005	
	P	Q	P	Q	P	Q	P	Q	P	Q
A	2	10	2	12	3	15	4	20	4	18
B	5	12	6	10	6	5	7	6	8	5
C	10	3	11	2	12	3	12	1	13	2

a) Compruebe los índices de precios según los modelos de Laspeyres y Paasche.

b) Hallar los índices complejos no ponderados (media aritmética, geométrica, armónica, etc.).Utilice como base el año 2001 en todos los casos.

5.- El dueño de un apartamento tenía pactado en 2012 un alquiler con su habitante de 600 dólares mensuales. Si en 2015 quiere examinar el contrato en base a los incrementos del IPC en esos años, cuyos índices han sido:

AÑOS 2012 2013 2014 2015
I. P. C. 108.3 105.5 103.3 103.8

¿Cuál será el nuevo precio de alquiler en un período mensual?

6.- A partir de los siguientes datos:

AÑO	CARBÓN		HIERRO	
	P	Q	P	Q
2009	5	2	8	3
2010	4	3	6	4

Conociendo que el peso del carbón es el triple que el del hierro, obtener el valor del índice complejo ponderado de 2006 con respecto a 2005, para los precios y cantidades indicadas utilizando la media agregativa.

7.- Presumamos que la beca de un estudiante de postgrado en 2011 era de $2400 y en 2014 de 2650. Se conoce que el IPC ha variado de 122.9 a 133.4 en esos años. ¿Cómo se ha variado el poder adquisitivo del estudiante? ¿Cuál debería de haber sido el valor de la beca en el 2014 para que su nivel de vida no hubiera cambiado respecto al año 2011?

8.- La información sobre la producción y precios de venta de una compañía en los años 2001 y 2003 por ramas de actividad de la empresa son:

RAMA	PRODUCCIÓN A		PRECIOS	
	2001	2003	2001	2003
A	100	150	14	10
B	250	300	18	12

- Calcule los índices de Laspeyres y Paasche para los precios en 2003 con base 2001.

- ¿Cuál ha sido la tasa anual de variación de los precios en el período 2001-2003 y la tasa de variación de dicho período?

9.- El IPC aumentó un 1% entre 2002 y 2003 y un 2% entre 2001 y 2002.

a) Consiga la serie de índices con año base 2001.

b) El valor de un determinado producto en el año 2002 era de $180. Conociendo que el valor de ese producto ha incrementado un 15% en procesos corrientes entre 2002 y 2003, obtenga

- el valor para el año 2003.
- el porcentaje de incremento o disminución del valor de este producto.

10.- En la siguiente tabla se presenta la información relacionada a los precios medios y los consumos de las dos tipos de hidrocarburantes para vehículos más usados en el país:

Año	Gasolina Súper		Gasolina Extra	
	Precio $	Consumo gal.	Precio $	Consumo gal.
2006	1,03	7,12	0,95	35,73
2008	1,12	6,56	1,14	36,19

- Realice un índice de tipo Laspeyres, ¿cuál ha sido la variación de los precios de los combustibles en el período 2006-2008?
- Calcule la participación de cada tipo de combustible en la variación del índice de precios de Laspeyres en el período 2006-2008.

11.- El PIB de los factores de producción en el Ecuador del 2005 ascendió a 2,4 billones corrientes y en 2010 ascendió a 2,4 billones. ¿Cuáles fueron los valores de dichos PIB en $ constantes de 2005 si el IPC fue 100 y 719 respectivamente?

12.- Los valores del patrimonio, en millones de dólares, de una cierta empresa, son:

AÑOS	VALORES
2005	43
2006	47
2007	52
2008	54

Si presumimos un 5% de inflación cada año respecto a que uno precede del otro, calcular los índices deflactados, para ello el utilice el índice adecuado.

13.- Una empresa fabricadora de impresoras en 3D fijó el precio unitario, en $900 en el año 2011. Por otro parte, el IPC del grupo de los mismos artículos al que pertenece dicho bien tomó los siguientes valores en el periodo 2011-2015:

AÑOS	VALORES
2011	800
2012	880
2013	920
2014	1000
2015	1060

a) Calcular el precio unitario del bien en cada uno de los años del lapso 2011-2015, si se presume que alcanzó igual progreso que el grupo de precios de los tales bienes.

b) Si el valor de las ventas de dicho bien en el año 2015 fue de $ 1.100.250, determinar dicho valor a precios constantes del año 2011.

4.9 Ecuaciones introducidas en el Capítulo 4

Números Índices

Un índice o número índice calcula el cambio en un concepto en particular (un producto o servicio) entre dos períodos, es el número que expresa el cambio referente en precio, cantidad o valor comparado con un período base.

[4.1] $$I = \frac{I(n+1)}{I(n)} * 100$$

Índices sin ponderar: Este número índice es la forma sencilla perteneciente al índice de agregados no ponderados.

[4.2] $$I(x) = \frac{\sum Q_I}{\sum Q_0} * 100$$

Índices ponderados: Cada dimensión o componente tiene un valor diferente asignado en función de diversos criterios.

[4.3] $$I_P = \frac{\sum P_I Q}{\sum P_0 Q} * 100$$

Método de Laspeyres

Conocida como la media aritmética ponderada de índices simples de cada concepto usándose como ponderación para cada bien.

[4.4] $$L_P = \frac{\sum P_I Q_0}{\sum P_0 Q_0} * 100$$

Método de Paasche

Representa la media aritmética ponderada de los números índices simples de cada concepto usándose como ponderación para cada artículo.

[4.5] $$P_P = \frac{\sum P_I Q_i}{\sum P_0 Q_i} * 100$$

Método de agregados con peso fijo

Este método posee una técnica para determinar pesos a los elementos de de estudio

[4.6] $$F_P = \frac{\sum P_i Q_2}{\sum P_0 Q_2} * 100$$

Método de promedio no ponderado relativos

Cuando tenemos más de un objeto (o actividad), primero hallamos el cociente del precio vigente entre el precio base para cada objeto y multiplicamos cada cociente conseguido por 100.

[4.7] $$x_R = \frac{\sum(\frac{P_i}{P_0} * 100)}{n}$$

Método de promedio ponderado relativos

Los métodos de promedio pesado de relativos permiten de varias maneras, establecer un valor con peso.

[4.8] $$x_R = \frac{\sum(\frac{P_i}{P_0} * 100)(P_n Q_n)}{\sum P_n Q_n}$$

Unidad 5

TEORÍA DE CONJUNTOS, COORDINATORIA Y DE PROBABILIDADES

“La probabilidad estadística de que las estructuras orgánicas y las reacciones tan precisamente armonizadas que caracterizan a los organismos vivos se generen por accidente, es cero. Así es, cero”.

-YLYA PRIGOYINE-

Físico, ingeniero químico, sistémico y profesor belga de origen ruso, galardonado con el Premio Nobel de Química (1977) y autor de "Estructura, estabilidad y fluctuaciones" (1971) "Orden en el caos" (1984), "Las leyes del caos" (1992), "La nueva alianza" (1983) y "El fin de las certidumbres" (1997).

CONTENIDO DE LA UNIDAD

INTRODUCCIÓN

En este apartado se muestra el diseño y estructura de un método de aprendizaje de la teoría de conjuntos, teniendo en cuenta, que en los objetos seleccionados influye el orden en que se toman, sean independientemente uno de otro, o ambos valores a la vez, así como el planteamiento de probabilidades en un nivel técnico; se mostrará las operaciones de álgebra de conjunto; para así estructurar el desarrollo de procesos aleatorios tomando en consideración las reglas de probabilidad.

OBJETIVOS DE APRENDIZAJE

Al finalizar el capítulo, será capaz de:

- Definir los términos relacionados a las operaciones de conjunto.
- Distinguir principales operaciones algebraicas de la teoría de conjuntos, análisis combinatorios y probabilísticos.
- Establecer los criterios para el uso y empleo de datos probabilístico.
- Describir las características de los conjuntos y su aplicación en el cálculo de probabilidades.
- Describir las características de la distribución de probabilidad

El siguiente capítulo es utilizado para la representación de la estadística,para determinar la incidencia que posee un evento ante la realización de ensayos o conocidos métodos probabilísticos. La información que se obtiene mediante el uso de tales métodos probabilísticos a menudo interviene en la toma de decisiones debido a que genera un criterio posible o imposible, así como en la resolución de problemas en diversas áreas de la ingeniería como los negocios y científicas médicas.

5.1 Conjuntos

5.1.1 Definición de conjunto

Un conjunto es la colección, reunión o agrupación de objetos que tienen una característica y están perfectamente definidos.

La descripción de un conjunto se puede identificar de diferentes maneras:

Por ***Comprensión,*** es cuando se describen las características que conforman al conjunto. Ej.

A = {las vocales}

B = {días de la semana}

C = {regiones del Ecuador}

Por ***Extensión o Tabulación***, cuando se enlistan todos los elementos que contienen al conjunto. Ej.

A = {a, e, i, o, u}

B = {Lunes, Martes, Miércoles, Jueves, Viernes, Sábado, Domingo}

C = {Costa, Sierra, Oriente, Insular}

Extensión por determinación, Si tenemos el conjunto:

A = {1, 2, 3, 4, 5, 6}

Para convertir en extensión por determinación se representa:

$$A = \left\{ {}^{X}/_{X} \in \mathcal{R}; 1 < x < 6 \right\}$$

Para describir los elementos de la respuesta:

$$A = \left\{ X \,/\, \underbrace{X \in \mathcal{R}; 1 < x < 6}_{\text{Condiciones}} \right\}$$

Tal que

Forma del Elemento

Lo que significa que para obtener el elemento ***X***, se debe cumplir con las dos condiciones:

De que ***X***, pertenezca a los números reales $(\boldsymbol{X} \in \boldsymbol{\mathcal{R}})$.

Que, ***X*** este entre **1** y **6**. $(\mathbf{1} < X < 6)$.

Por lo tanto, la respuesta se lee como ***X*** tal que ***X*** pertenece a los números reales, donde ***X*** se encuentre entre 1 y 6.

Por ***Diagramas de Venn***, que es la representación gráfica de los elementos que conforman el conjunto. Ej.

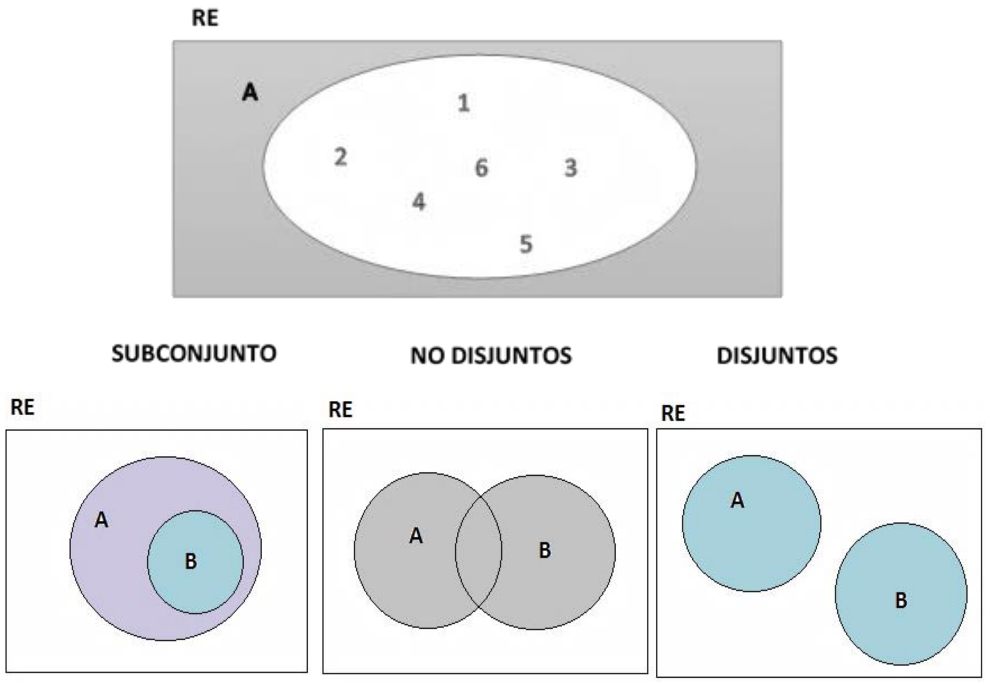


Ilustración 5.1 Diagramas de Venn

5.1.2 Simbología de Conjunto

Tabla 5.1

Principales simbologías

Simbología de Conjuntos	
Símbolo	**Descripción**
{ }	Conjunto
$\in$	Es un elemento del conjunto o pertenece al conjunto.
$\notin$	No es un elemento del conjunto o no pertenece al conjunto.
/	Tal que.
n (C)	Cardinalidad del conjunto C.
U ; RE	Conjunto universo o referencial.
Φ	Conjunto vacío.
$\subseteq$	Subconjunto de.
$\subset$	Subconjunto propio de.
$\not\subset$	No es subconjunto propio de.
$>$	Mayor que.
$<$	Menor que.
$\geq$	Mayor o igual que.
$\leq$	Menor o igual que.
$\cap$	Intersección de conjuntos.
$\cup$	Unión de conjuntos.
A^c	Complemento del conjunto A.
$=$	Símbolo de igualdad.
$\neq$	No es igual a.
…	El conjunto continúa.
$\Leftrightarrow$	Si y sólo sí.
$\neg$	No, negación lógica (es falso que).
$\wedge$	Y
$\vee$	O
$\equiv$	Es equivalente
$\mathcal{R}$	Números reales

5.1.3 Clasificación de los conjuntos

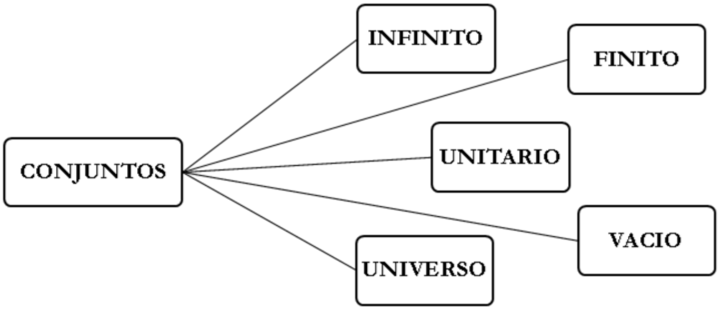


Ilustración 5.2 Principales Conjuntos

Conjunto infinito: no tiene cantidad finita de elementos. Ej.

A = {números pares positivos}

Conjunto finito: tiene exactamente elementos finitos con principio y fin. Ej.

A = {habitantes de la ciudad}

Conjunto universal: es el que se encuentra conformado por elementos del tema referencial.

U; RE = {letras del alfabeto español}

Conjunto vacío: es el que se encuentra conformado por elementos del tema referencial.

A = {número par e impar a la vez} → { } = Φ

Conjunto unitario: se encuentra conformado por un elemento determinado del tema referencial.

A = {1}

5.1.4 Principio de extensión

Dos conjuntos son iguales si y solamente si tienen los mismos elementos. Es decir que si cada elemento que pertenece al conjunto A pertenece también al conjunto B y si cada elemento que pertenece al conjunto B pertenece también al conjunto A.

$$A = \{1, 2, 3, 4, 5\} \qquad A = B \qquad B = \{2, 4, 5, 1, 2\}$$

Un conjunto no cambia si se reordenan los elementos.

$$C = \{5, 6, 7, 8, 9\} \qquad C = D \qquad D = \{9, 8, 7, 6, 5\}$$

Un conjunto no cambia si repiten algunos de sus elementos.

$$E = \left\{ X /_{X^2 - 3X = -2} \right\} \qquad F = \{2; 1\} \qquad G = \{1; 2; 2; 1\}$$

$$X^2 - 3X + 2 = 0$$

$$(X - 2)(X - 1) = 0$$

$$X = 1; X = 2$$

5.1.5 Subconjunto

Si cada elemento de un conjunto A es también elemento del conjunto B entonces B es un conjunto de A.

$A \subset B$: Se lee el subconjunto B está contenido en A.

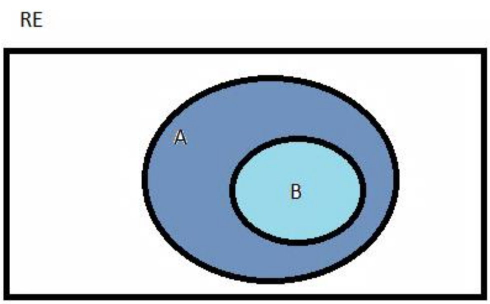


Ilustración 5.3 Gráfico de un subconjunto

5.1.6 Teorema de los subconjuntos

1.- Para todo conjunto se entiende que el conjunto vacío es un subconjunto del conjunto universo.

$$\Phi \subset A \subset U$$

2.- Para todo conjunto A se entiende que A, es un conjunto de A, por lo tanto cada conjunto es un subconjunto propio de sí mismo.

$$A \subset A$$

3.- Si A es un conjunto de B y B es un conjunto de C, entonces A es un conjunto de C.

$$(A \subset B) \wedge (B \subset C) \Rightarrow (A \subset C)$$

4.- Si A es igual que B, si y solamente si A es un subconjunto de B y B es un subconjunto de A.

$$(A \subset B) \equiv (A \subset B) \wedge (B \not\subset A)$$

5.1.7 Conjunto complemento

El complemento de un conjunto A es el conjunto de elementos que pertenecen a **U** (Universal), pero que no pertenece a otro conjunto A.

$$\mathbf{A^c} = \left\{ \boldsymbol{X} / (\boldsymbol{X} \in \mathbf{U}) \wedge (\mathbf{x} \notin \mathbf{A}) \right\}$$

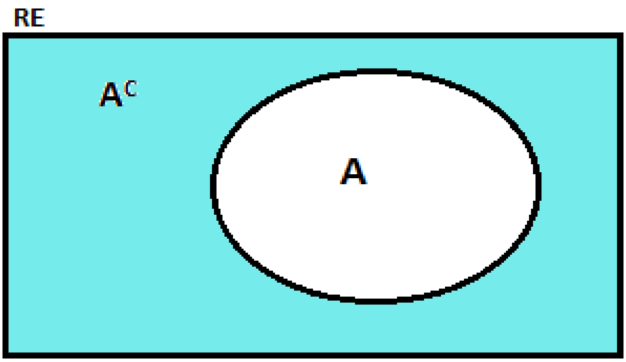


Ilustración 5.4 Gráfico de un conjunto complemento

5.1.8 Conjunto propio

Cuando A es un subconjunto propio del conjunto B **(A ⊂ B);** si todo elemento de A es también elemento de B y además existe un elemento de B que no es elemento de A.

$$(A \subset B) \equiv (A \subset B) \wedge (B \not\subset A)$$

Ejemplo 5.1: Ya que todo conjunto de A es subconjunto de sí mismo, se dice que B es subconjunto propio de A.

$$\mathbf{U} = \left\{ \boldsymbol{X} / (\boldsymbol{X} \in \boldsymbol{\mathcal{R}}); (\mathbf{1} < X < 10) \wedge (\mathbf{10} \notin \mathbf{A}) \right\}$$

A= {1, 2, 3, 4, 5, 6, 7, 8, 9} B= {1, 3, 5, 7, 9} **A ⊂ B**

$A^C = \{10\}$

5.1.9 Conjunto potencia

Es la familia de todos los subconjuntos de un subconjunto. Se representa como norma:

$$\boldsymbol{P(A) = 2^n}$$

Donde:

P(A) representa al conjunto potencia.

n es el número de elementos del conjunto.

Ejemplo 5.2: Dado el conjunto potencia **A = {0, 1}**; se generan los subconjuntos:

A = {(0, 1); (0); (1); Φ}

Que representan a 4 subconjuntos, donde la cardinalidad está dada por el número de elementos (0, 1) que son 2 elementos.

$P(A) = 2^2 = 4 \rightarrow 4$ subconjuntos

Ejemplo 5.3: Para el conjunto **B = {0, 1, 2};** se generan los subconjuntos:

B **= ({0, 1, 2}; {0}; {1}; {2}; {0, 1}; {0,2}; {1, 2}; {})**

$P(A) = 2^3 = 8$ subconjuntos

5.2 Operaciones de conjuntos

5.2.1 Unión de conjuntos

Unión de dos conjuntos A y B, es el conjunto formado por todos los elementos que pertenecen al conjunto A o al B, o ambos.

$$(\mathbf{A} \cup \mathbf{B}) = \left\{ \mathbf{X} / (\mathbf{X} \in \mathbf{A}) \cup (\mathbf{X} \in \mathbf{B}) \right\}$$

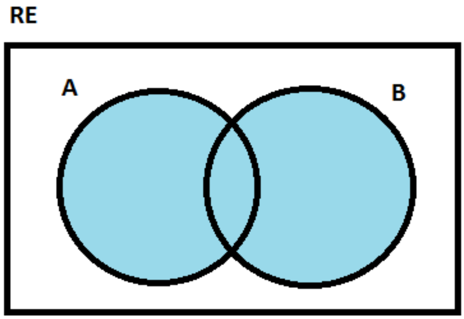


Ilustración 5.5 Gráfico de unión de conjuntos

Ejemplo 5.4: Mediante los siguientes conjuntos (A, B, C): A= {1, 2, 3, 4, 5}; B= {3, 4, 5, 6, 7}; C= {2, 3, 5, 7}.
Encontrar los siguientes conjuntos (A U B); (A U C); (B U C):

- A U B = {1, 2, 3, 4, 5, 6, 7}
- A U C = {1, 2, 3, 4, 5, 7}
- B U C = {2, 3, 4, 5, 6, 7}

5.2.2 Intersección de conjuntos

La intersección de los conjuntos A y B, es el conjunto formado por los elementos que pertenecen tanto al conjunto A como al conjunto B simultáneamente.

$$(\mathbf{A} \cap \mathbf{B}) = \left\{\mathbf{X} / _{(\mathbf{X} \in \mathbf{A}) \cap (\mathbf{X} \in \mathbf{B})}\right\}$$

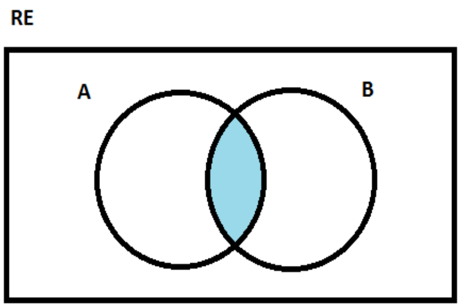


Ilustración 5.6 Gráfico de intersección de conjuntos

Ejemplo 5.5: Mediante los siguientes conjuntos (A, B, C): A= {1, 2, 3, 4, 5}; B= {3, 4, 5, 6, 7}; C= {2, 3, 5, 7}.

Encontrar los siguientes conjuntos (A ∩ B); (B ∩ C), (A ∩ C).

- $A \cap B = \{3, 4, 5\}$
- $B \cap C = \{3, 5, 7\}$
- $A \cap C = \{2, 3, 5\}$

Ejemplo 5.6: Sea los conjuntos:

H = {estudiantes varones de un curso}; M = {estudiantes mujeres del mismo curso}.

Hallar:

- $H \cap M$
- H U M.
- H U M = {Curso de Estudiantes};
- $H \cap M = \{\Phi\}$

5.2.3 Diferencia de conjuntos

La diferencia de conjuntos A y B es el conjunto formado por todos los elementos de A que no pertenecen a B.

$$(\mathbf{A} - \mathbf{B}) = \left\{ \mathbf{X} / (\mathbf{X} \in \mathbf{A}) \wedge (\mathbf{X} \notin \mathbf{B}) \right\}$$

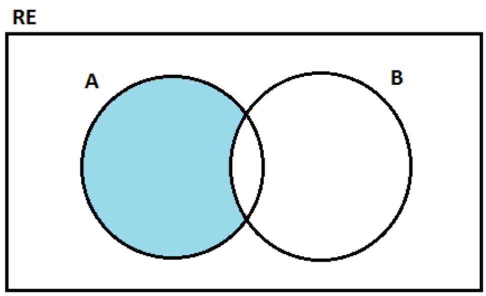


Ilustración 5.7 Gráfico de diferencia de conjuntos

Ejemplo 5.7: Si los conjuntos A y B están representados por: A = {a, b, c, d, e} y B = { a, e, i, o }, entonces la diferencia de conjuntos será representada por todos los elementos que estén en A, esto es:

A – B = {b, c, d}

Ejemplo 5.8: Sean los conjuntos:

- $\mathbf{W} = \left\{ {}^{\mathrm{X}}\!/_{\mathrm{X}} \; ; (número\ impar) \wedge (\mathrm{x} < 10) \right\}$
- $\mathbf{Z} = \{6,7,8,9,10,11,12\}$

Entonces, que elementos conforman el conjunto **W – Z.**

$$\mathbf{W} - \mathbf{Z} = \{1,3,5\}$$

5.2.4 Propiedades de las operaciones entre conjuntos

Tabla 5.2

Operaciones de los conjuntos

Propiedad	Unión	Intersección
Asociativa	(A U B) U C = A U (B U C)	(A ∩ B) ∩ C = A ∩ (B ∩ C)
Conmutativa	(A U B) = (B U A)	(A ∩ B) = (B ∩ A)
Idempotencia	A U A = A	A ∩ A = A
Absorción	A U (B ∩ A)= A	A ∩ (B U A)= A
Distributiva	A U (B ∩ C)= (A U B) ∩ (B U A)	A ∩ (B U C)= (A ∩ B) U (B ∩ A)
Neutralidad	A U Ø = A	A ∩ U = A
	A U U = U	A ∩ Ø = Ø
Complementación	A U A^c = U	A ∩ A^c = Ø
Ley de Morgan	$(A U B)^c$ = A^c ∩ B^c	$(A ∩ B)^c$ = A^c U B^c
Reducción	A U (B ∩ A) = A	A ∩ (B U A) = A
Identidad	A U Ø = A	A ∩ U = A
Absorción	A U Re = Re	A ∩ Ø = Ø
Involución	$(A^c)^c$ = A	
Diferencia	A - B = $(A ∩ B)^c$	
Complemento	U^c = 0 Ø = RE	

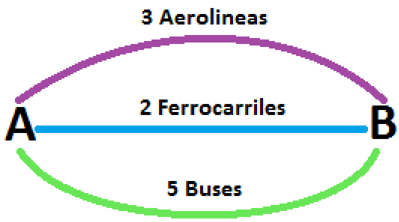
5.2.5 Principios fundamentales de conteo

Para empezar, previamente identificamos los principios fundamentales de conteo, los cuales se describen a continuación:

La adición o suma

Es la operación matemática en la que consiste reunir cantidades de tal manera que de origen a un resultado, donde el orden de los elementos no altera el resultado.

Ejemplo 5.9: De cuántas maneras posibles se puede ir del punto A al punto B.



$$3 + 2 + 5 = 10$$

Existen 10 maneras distintas de llegar del punto A hacia B.

La sumatoria:

La sumatoria consiente sintetizar la notación de una suma cuyos términos permiten cierta ley de formación.

Ejemplo 5.10: Para indicar la sumatoria de: $\sum_{x=1}^{7} ax$

Se describe como sumatoria de a_x, donde la variable x se reemplaza por los valores del 1 a 7.

$$ax = a(1) + a(2) + a(3) + a(4) + a(5) + a(6) + a(7)$$

$$ax = a(1+2+3+4+5+6+7) \text{ Factor común}$$

$$ax = a(28)$$

La sumatoria de la variable x que tiende de 1 a 7, es igual a 28, multiplicado con la variable *a* da como resultado 28 *a.*

Ejemplo 5.11: Obtener la sumatoria de: $\sum_{x=1}^{4} x$

$$x = (1)+ (2)+ (3)+ (4)$$

$$x = 9$$

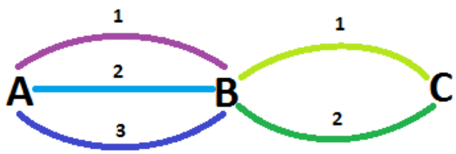
La sumatoria de la variable x que tiende de 1 a 4, es igual a 9.

La multiplicación

Se da cuando los sucesos o eventos, ocurren el uno a continuación del otro, son como un suceso compuesto.

[5.1] $n_1 * n_2 * n_3 * \ldots * n_x$

Ejemplo 5.11: Hallar el número de rutas para ir de A hacia C.



Del punto A hacia el B existen 3 rutas y del punto B al C, hay 2 rutas:

$$3 * 2 = 6$$

En conclusión existen 6 maneras de llegar del punto A hacia el C.

Ejemplo 5.12: ¿Cuántos puntos muestrales hay al momento de lanzar un par de dados?

$n_1 * n_2 = (6)*(6) = 36$ formas posibles.

Factorial de un número
La factorial de un número natural *n* mayor que uno (1) es igual al producto de los *n* principales números naturales; el símbolo característico es "!". Así:

[5.2] $$n! = n!\ \forall\ (n \in R)$$

De la definición teoriza que la factorial de un número es similar al producto de dicho número por la factorial del
anterior.

En general: n (n[1])= n!

Además se define: 0! = 1 y 1! = 1

Ejemplo 5.13: Desarrollar el siguiente número factorial.

$$6! = 6*5*4*3*2*1! = 720$$

Ejemplo 5.14: La factorial de la operación:

$$n! = \frac{13!}{9!}$$

$$n! = \frac{13!}{9!} = \frac{13*12*11*10*\cancel{9!}}{\cancel{9!}} = 17{,}160$$

5.3 Análisis combinatorio

5.3.1 Definición de análisis combinatorio
El análisis combinatorio o combinatoria es parte de las matemáticas que trabaja los diferentes procedimientos en que se pueden integrar agrupaciones entre elementos de uno o más conjuntos y como contar ordenadamente su número.

Ejemplo 5.15: Un empresa fabrica camisas en 5 estilos diferentes y 4 colores distintos para cada una, si la empresa desea mostrar a su clientela los nuevos modelos de camisas de todos los estilos y colores. ¿Cuántos modelos diferentes deberían colocar en el aparador?

$n_1 = 5$
$n_2 = 4$ $\qquad 5 * 4 = 20$

Se colocarían 20 modelos diferentes en el aparador

Ejemplo 5.16: En Ecuador, los vehículos tienen una matrícula que consta de tres letras sacadas de un alfabeto de 26, seguidos de cuatro dígitos decimales. ¿Cuántas matrículas distintas pueden llegar a haber?

$(10^4)*(26^3) = 10 * 10 * 10 * 10 * 26 * 26 * 26 = 175.760.000$

Existen 175.760.000 posibles matrículas.

Ejemplo 5.17: Si una prueba de selección múltiple consta de 5 preguntas, para cada una con 4 respuestas posibles de la prueba y solo una es correcta.

En cuántas formas diferentes puede un estudiante escoger una respuesta para cada una.

a.- $4^5 = 1024$

En cuántas formas puede un estudiante escoger la alternativa para cada una y tener las respuestas correctas.

b.- $3^5 = 243$

Ejemplo 5.18: Cuatro parejas compran 8 boletos para ir al cine, se desea saber de cuántas formas posibles, pueden sentarse las parejas bajo los siguientes criterios:

Sentarse sin restricciones.

Como deciden sentarse sin restricciones nos encontramos ante una permutación debido a que se va a ordenar los elementos.

${}_8P_8 = 8! = 8 * 7 * 6 * 5 *4 *3 *2 * 1! = 40.320$

Hay 40.320 números de formas posibles que se pueden ordenar las personas.

Sentarse por parejas.

Tenemos un análisis combinatorio y permutación, la permutación es porque se ordenan por parejas, lo que sería:
${}_2P_2 = 2!$

Se pueden agrupar las parejas de 4 formas posibles

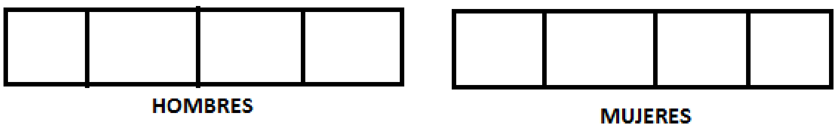
${}_4P_4 = 4!$

El análisis combinatorio ocurre a continuación como un suceso compuesto, por lo tanto corresponde a una multiplicación, siendo el resultado final.

$4! * 2! = 48$

Si todos los hombres deciden sentarse juntos a la derecha de todas las mujeres.

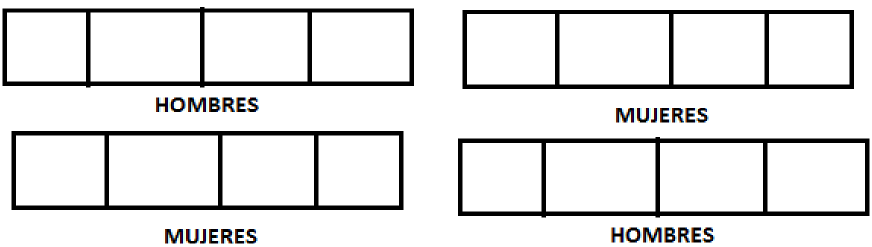
Tenemos dos bloques de permutaciones que corresponden a un análisis combinatorio de multiplicación.



$$4! * 4! = 576$$

A. Si todos los hombres deciden sentarse juntos y las mujeres también.

A diferencia del literal C pueden sentarse en el orden que quieran, derecha o izquierda por lo que el análisis combinatorio corresponderá:



$$2! * 4! * 4! = 1.152$$

5.4 Permutaciones

5.4.1 Definición de permutación

La permutación es el arreglo de todos los objetos del conjunto o parte de estos, a la permutación le importa el orden de los elementos involucrados.

[5.3] $${}_{n}P_{r} = n!$$

5.4.2 Características de las permutaciones

- Selecciona todos los elementos dados para realizar el análisis.
- Se toma importancia al orden de los elementos.
- No se repiten los sucesos.

Permutación sin repetición

Si en el arreglo del orden de los elementos de la permutación corresponda a una parte de ellos, se representará mediante la fórmula:

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

Donde:

n: Corresponde a todas los elementos.

r: Número de elementos deseados.

El número de permutaciones de ***n*** elementos distintos es ***(n - 1)!***.

Ejemplo 5.19: Se sacan dos billetes de la lotería de 30 posibles encuentren cuantas formas puede escoger dos billetes para el primero y segundo premio.

n = 30 (billetes)

r = 2 (premios)

$$_{n}\mathbf{P}_{r} = {}_{30}\mathbf{P}_{2} = \frac{30!}{(30-2)!} = \frac{30*29*28!}{28!} = 870$$

Constan 870 posibles maneras de escoger los billetes ganadores

Ejemplo 5.20: De cuántas formas posibles se puede disponer a tres conferencistas en 3 congresos diferentes si ellos están disponibles en 5 fechas.

n = 5 (días de congreso)

r = 3 (conferencistas)

$$_{5}P_{2} = \frac{5!}{(5-3)!} = \frac{5*4*3*\cancel{2!}}{\cancel{2!}} = 60$$

Es posible disponer 60 conferencias para los conferencistas

Permutación con repetición

Se llama permutación con repetición, al número de permutaciones de distintos objetos en la que posee elementos repetidos de primer orden, de un segundo orden, hasta de enésimo orden y que este cumpla el orden de los eventos

.

[5.5] $$_{n}\mathbf{P}_{r1,\, r2,\, r3} = \frac{\boldsymbol{n}!}{\boldsymbol{r_1}!+\boldsymbol{r_2}!+\cdots+\boldsymbol{r_k}!}$$

Ejemplo 5.21: En cuántas maneras diferentes pueden acomodarse 3 reflectores rojos, 4 amarillos y 3 verdes para alumbrar un sector de la ciudad.

.n = 10 (total de elementos)

$\mathbf{r_1}$ = 3 (reflectores rojos)

$\mathbf{r_2}$ = 4 (reflectores amarillos)

$\mathbf{r_3}$ = 3 (reflectores verdes)

$$_{10}P_{r3,\, r4,\, r3} = \frac{10!}{3!4!3!} = \frac{3628800}{864} = 4200$$

Existen un total de 4200 maneras de acomodar los 3 reflectores rojos, 4 amarillos y 3 verdes para alumbrar el sector.

Ejemplo 5.22: De cuántas formas se pueden colocar 8 personas en dos habitaciones triples y una habitación doble.

n = 8 (personas)

$\mathbf{r_1}$ = 3 (habitación triple)

$\mathbf{r_2}$ = 3 (habitación triple)

$\mathbf{r_3}$ = 2 (habitación doble)

$${}_{8}P_{r3r3r2} = \frac{8!}{3!3!2!} = 560$$

Constan 560 posibles maneras de ubicar a las 8 personas en las habitaciones.

5.5 Combinaciones

5.5.1 Definición de combinaciones

Para las combinaciones son importantes las formas posibles de seleccionar $\boldsymbol{r}$ objetos de un total de $\boldsymbol{n}$, sin importar el orden. El número de combinaciones de $\boldsymbol{n}$ objetos distintos, tomando $\boldsymbol{r}$ a la vez.

[5.6] $${}_{n}C_{r} = \frac{\boldsymbol{n}!}{\boldsymbol{r}!(\boldsymbol{n}-\boldsymbol{r})!}$$

Donde:

n: Corresponde a los elementos.

r: Número de elementos deseados, sin orden.

5.5.2 Características de las combinaciones

.

- No selecciona o usa todos los elementos dados para realizar el análisis.
- No importa el orden de los elementos.
- Los sucesos se diferencian entre sí, ya que uno posee elementos que el otro no tiene.

Ejemplo 5.23: Un gremio sindical de 35 trabajadores quiere formar una delegación de 3 empleados. ¿Cuántas delegaciones pueden haber?

n = 35 (trabajadores)

r = 3 (integrantes de la delegación)

$$ {}_{35}C_{3} = \frac{3!}{3!(35-3)!} = \frac{35*34*33*\cancel{32!}}{3!\cancel{(32!)}} = 6545 $$

Ejemplo 5.24: Encuentre el número de comités que puede formarse con 4 ingenieros químicos y 3 ingenieros físicos que comprenden 2 químicos y un físico.

$$ {}_{4}C_{2} = \frac{4!}{2!(4-2)!} = 6 $$

$$ {}_{3}C_{1} = \frac{3!}{1!(3-1)!} = 3 $$

6 * 3 = 18

5.6 Probabilidades

5.6.1 Definición de probabilidades

La etimología o término proviene del adjetivo latino *probabilis* que originalmente significaba digno de aprobación, "loable", fue como equivalente del griego de "persuasivo", "creíble", "verosímil", o sea, "probable". Una simple traducción al predicado con estas indicaciones daría como resultado que la probabilidad es la aseveración de un criterio que puede parecernos probable o verosímil en un mismo grado o un grado mayor o menor que otra.

Dos aplicaciones primordiales de la teoría de la probabilidad aplicadas en el día a día son, en el análisis de riesgo de la bolsa de valores y en el comercio de los mercados de materias primas (importaciones y exportaciones). Los gobiernos regularmente emplean métodos probabilísticos en medidas económicas y a menudo calculan la prosperidad utilizando métodos que son basados en la utilización de variables aleatorias.

Las Probabilidades corresponden a la rama de la matemática que realiza ciertos experimentos llamados aleatorios, o sea que regidos por el azar, en que se está al tanto de todos los resultados posibles, pero no es posible tener certeza de cuál será en particular la consecuencia del experimento.

Por ejemplo, los experimentos aleatorios usuales son el lanzamiento de una moneda, el lanzamiento de un dado, extracción de una carta de un mazo de naipes.

5.6.2 Teoremas de probabilidad

A.- La probabilidad es determinable, esto quiere decir que existe la posibilidad entre 0 y 1.

.

- P (x) = 1; Evento Cierto.
- P (x)= 0; Evento Incierto.

Ejemplo 5.25: Eventos ciertos e inciertos.

a) P (sol) = salga mañana = 1
b) P (pasar sin estudiar) = 0

B.- Según el espacio muestral se clasifican en finito e infinito.

- Finito: se conoce el número de eventos posibles.
- Infinitos: se desconoce el número de eventos posibles.

Ejemplo 5.26: Espacio finito o infinito.

a) El lanzamiento de una moneda.
b) El lanzamiento de un dado.
c) Extracción de una carta de un mazo de naipes.

C.- Según la forma de aplicarse o enfocarse, a probabilidad se clasifica en:

- **Modelo de frecuencia relativa:** se da por la observación empírica (simple), registra la frecuencia con la que ocurrió un evento en el pasado y se estima en el futuro próximo.
- **Modelo Subjetivo:** consiste en asignar una probabilidad a un evento nunca ocurrido.
- **Modelo clásico:** determina la probabilidad antes de ser observado.

Ejemplo 5.27: Modelo de frecuencia relativa

En cierto año hubo 500 nacimientos en un hospital, de los cuales 320 eran niñas. ¿Cuál es la probabilidad que el siguiente año nazcan varones?

$$\frac{500-320}{500}=0{,}36$$

Ejemplo 5.28: Modelo subjetivo

La probabilidad de que alguien de la clase se saque la lotería.

La probabilidad que nevará en el desierto.

Ejemplo 5.29: Modelo clásico

Lanzar un dado, para que salga 2.

$$P(2)=\frac{1}{6}$$

Escoger una J, de un mazo de naipes.

$$P(4) = \frac{4}{52}$$

5.6.3 Conceptos básicos de probabilidad

Experimento. Un experimento aleatorio es un sistema de procesos que tiene las siguientes propiedades:

- El proceso se perpetra de acuerdo a un conjunto bien definido de reglas.
- Son naturaleza tal que se repite o puede idear la repetición del mismo.
- El efecto de cada ejecución depende de "la casualidad" y por lo tanto, no se puede pronosticar un resultado único.
- Una sola realización del experimento se llama ensayo

Ejemplo:

- El lanzamiento de una moneda.
- El lanzamiento de un dado.
- Extracción de una carta de un mazo de naipes.

Espacio muestral. Conjunto de todos los posibles resultados de un experimento, se llama espacio muestra o espacio muestral del experimento, y se denota por S.

- Cada uno de los resultados del experimento se llama elemento o punto de S.
- Se dice que un espacio muestra es finito o infinito, cuando el conjunto S tiene un número finito o infinito de elementos, respectivamente.

Ejemplo:

Moneda = {Cara, Cruz}
Dado = {1, 2, 3, 4, 5, 6}

Evento. En varios casos se lo denomina evento o suceso a todo subconjunto de un espacio muestral.

Ejemplo:

A= {Obtener un numero par al momento de lanzar un dado}

A= {2, 4, 6}

Eventos mutuamente excluyentes. Dos eventos son mutuamente excluyentes si no pueden suceder en forma simultánea, esto es si y sólo si, su intersección es vacía.

Ejemplo 5.30: En el lanzamiento de un dado los eventos B = {2} y C = {5, 6} son mutuamente excluyentes por cuanto $B \cap C = \emptyset$

Eventos complementarios. Si $A \cap B = \emptyset$ y $A \cup B = E$, se dice que A y B son eventos complementarios. $A^c = B$ y $B^c = A$

Eventos independientes. Cuando A y B son dos eventos con probabilidades, entonces B es independiente de A, por lo tanto A es independiente de B. $(A \neq B) \wedge (B \neq A)$

5.6.4 Teorema de las probabilidades

La probabilidad de un evento A es la suma de los pesos de todos los puntos muestrales en A, por tanto:

$$0 \leq P(A) \leq 1 \quad P(\emptyset) = 0 \quad P(S) = 1$$

Si un experimento consigue obtener como resultado cualquiera de ***N*** diferentes resultados igualmente probables, y si exactamente en ***n*** de estos resultados comprenden al evento ***A***, entonces la probabilidad del evento es:

[5.7] $$P(A) = \frac{n}{N}$$

Ejemplo 5.31: Una funda de caramelos surtidos contiene seis mentas, cuatro masticables y tres bombones, si una persona selecciona uno al azar. ¿Qué probabilidad tiene de sacar a) menta b) masticable o bombón?

a) De los 13 dulces, 6 son mentas.

$$P(M) = \frac{6}{13} = 0.4615 \rightarrow 46\%$$

b) De los 13 dulces, 7 son masticables y bombones.

$$P(M) = \frac{7}{13} = 0.5384 \rightarrow 54\%$$

Hay una probabilidad que seleccione mentas del 46% y masticables con bombones un 54%.

Ejemplo 5.32: Una juego de Póker que consiste en 50 cartas, encuentre la probabilidad de tener cuatro ases.

$$P(M) = \frac{4}{50} = 0.08 \rightarrow 8\%$$

5.6.5 Reglas de la adición

Existen dos reglas de la adición, la regla especial de la adición y la regla general de la adición.

Regla especial de la adición: Para emplear la regla especial de la adición, los eventos obligatoriamente tienen que ser mutuamente excluyentes. Acordarse que mutuamente excluyentes significa que cuando un evento sucede, ninguno de los demás eventos puede pasar al mismo tiempo.

[5.8] **P (A U B) = P (A) + P (B)**

Ejemplo:

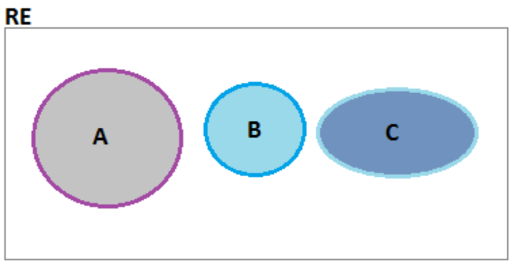


Ilustración 5.7 Gráfico de adición de conjuntos

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

Regla general de la adición: Los resultados de un experimento pueden no ser mutuamente excluyentes.

Cuando dos eventos ocurren al mismo tiempo, la probabilidad se denomina **probabilidad conjunta**.

[5.9] $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Ejemplo:

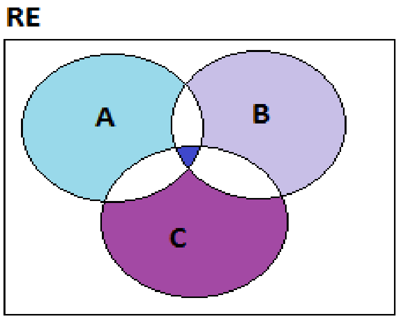


Ilustración 5.8 Adición de Conjuntos A, B, C

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cup B) - P(B \cup C) - P(A \cup C) + P(A \cap B \cap C)$$

Regla del complemento: Se pide para establecer la probabilidad de que un evento ocurra restando de 1 la probabilidad de un evento que no ha ocurrido.

[5.10] $$P(A) = 1 - P(\neg A) \quad \equiv \quad P(A) + P(\neg A) = 1$$

(Es equivalente)

Ejemplo:

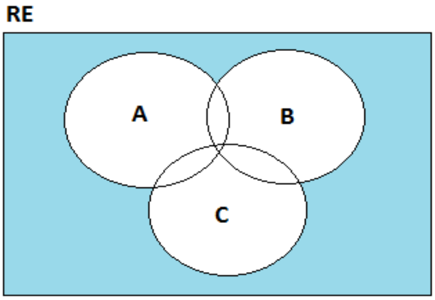


Ilustración 5.9 Conjunto Complemento

$$\neg P\ (A \cup B \cup C) = 1 - P\ (A \cup B \cup C)$$

5.6.6 Reglas de la multiplicación

La multiplicación estimará la probabilidad de que la ejecución de dos eventos sea simultánea. Hay dos reglas de la multiplicación, la regla especial y la regla general.

Regla especial de la multiplicación: La regla especial de la multiplicación pide que dos eventos, A y B, sean independientes, y lo son si el hecho de que uno suceda no altera la probabilidad de que el otro acontezca.

La **independencia** es un evento que acontece, no tiene ningún resultado sobre la probabilidad que algún otro evento acontezca.

[5.11] $$P\ (A \cap B) = P\ (A) * P\ (B)$$

Regla general de la multiplicación: Si dos eventos no son independientes, se dice que son dependientes, es decir sirve para determinar la probabilidad conjunta de dos eventos cuando éstos no son independientes.

La **dependencia** se da cuando al resultado de un experimento previo afecta al nuevo evento.

[5.12] $$P\ (A \cap B) = P\ (A) * P\ (B/A) \quad \equiv \quad P\ (B/A) = \frac{P\ (A \cap B)}{P\ (A)}$$

(Es equivalente)

Donde la designación apropiada de la fracción **B/A,** se lee como **"B"** dado **"A"**, por lo tanto es una **probabilidad condicional**, ya que su valor se encuentra condicionado (u obedece) por el hecho de que haya sido el primer evento en haber acontecido.

5.6.7 Teorema de Bayes

El teorema de Bayes es una situación probabilística en la es posible conocer las probabilidades de que acontezcan una serie de sucesos. En esencia, los seguidores de la estadística tradicional también denominada objetivista o frecuencialista, sólo admiten probabilidades apoyadas en experimentos repetibles y que tengan una validación empírica mientras que los llamados estadísticos bayesianos permiten y defienden la utilidad de las probabilidades subjetivas.

A partir de que ha ocurrido el suceso B (ha ocurrido un accidente) deducimos las probabilidades del suceso A (¿estaba lloviendo o hacía buen tiempo?).

La fórmula del Teorema de Bayes es:

[5.13] $$\mathbf{P\ (A\ /B) = \frac{P\ (\ A\ ){*}P\ (B/A)}{\sum[P(\ A\ ){*}P(B/A)]}}$$

Ejemplo 5.33: El departamento de meteorología de la Armada del Ecuador, ha anunciado tres posibilidades para el fin de semana, que llueva: probabilidad del 50%; que granice: probabilidad del 30%; que haya niebla: probabilidad del 20%.

Según estos posibles estados meteorológicos, la posibilidad de que ocurra un accidente es la siguiente:

- Si llueve: probabilidad de accidente del 10%.
- Si es graniza: probabilidad de accidente del 20%
- Si hay niebla: probabilidad de accidente del 5%.

a) Probabilidad de que llueva:

$$P\ (A\ /B) = \frac{(\ 0.50\ ) * \ (0.20)}{(0.50 * 0.20) + (0.30 * 0.10) + (0.20 * 0.05)} = 0.714 \rightarrow 71.4\%$$

La probabilidad de que estuviera lloviendo es del 71,4%.

b) Probabilidad de que granice:

$$P\ (A\ /B) = \frac{(\ 0.10\ ) * \ (0.10)}{(0.50 * 0.20) + (0.30 * 0.10) + (0.20 * 0.05)} = 0.214 \rightarrow 21.4\%$$

La probabilidad de que estuviera nevando es del 21,4%.

c) Probabilidad de que hubiera niebla:

$$P\ (A\ /B) = \frac{(\ 0.20\ ) * \ (0.05)}{(0.50 * 0.20) + (0.30 * 0.10) + (0.20 * 0.05)} = 0.071 \rightarrow 7.1\%$$

La probabilidad de que hubiera niebla es del 7,1%.

Ejemplo 5.34: Si se selecciona aleatoriamente a una persona de este grupo determine la probabilidad de que sea hombre dado que tiene educación de nivel secundaria.

a) De que no tenga título profesional dado que es mujer.
b) De que sea mujer dado que tiene instrucción secundaria.

$$P(H|S\ ) = \frac{P(H \cap B)}{P(S)} = \frac{28/200}{78/200} = \frac{0.14}{0.39} = 0.36$$

$$P(U'|M) = \frac{P(U' \cap M)}{P(M)} = \frac{95/200}{112/200} = 0.85$$

$$P(M|S\ ) = \frac{P(M \cap S)}{P(S)} = \frac{50/200}{78/200} = 0.64$$

Ejemplo 5.35: Tres joyeros tienen dos compartimientos de compras, el primer joyero hay un reloj de oro en compra. Del segundo joyero hay un reloj de plata, en el tercer joyero hay un compartimiento de un reloj de oro y uno de plata. Si seleccionamos un joyero aleatoriamente, abrimos unos de los compartimientos y hallaremos un reloj de plata. ¿Cuál es la probabilidad de que el otro sea un reloj de oro?

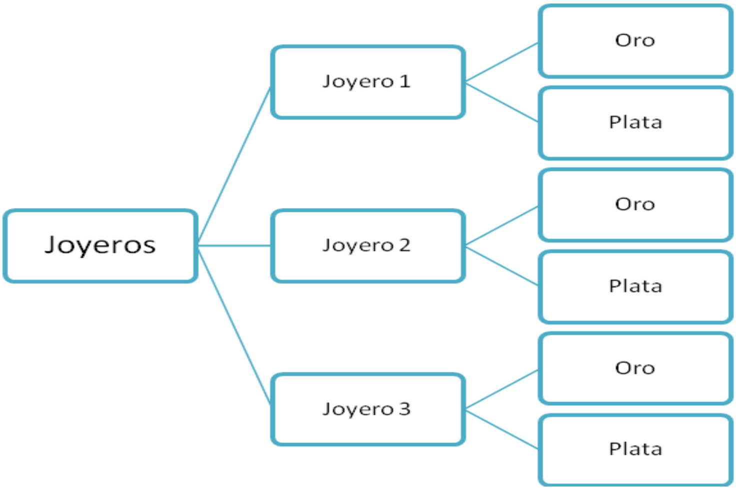


Ilustración 5.10: Gráfica del caso de los joyeros

$$P = \frac{1}{3} = 0.3333$$

La probabilidad de que el otro sea reloj de oro es de 33.33%

5.35.- Se sabe que un suero de la verdad, que se aplica a un sospechoso es 90% confiable cuando la persona es culpable y 99% cuando es inocente. En otras palabras 10% de los culpables se juzgan inocentes, por el uso de este suero y el 1% de los culpables. Se selecciona un individuo de un grupo de sospechosos de los cuales solo un 5% ha cometido un crimen y el suero indica que es culpable. ¿Cuál es la probabilidad de que sea inocente?

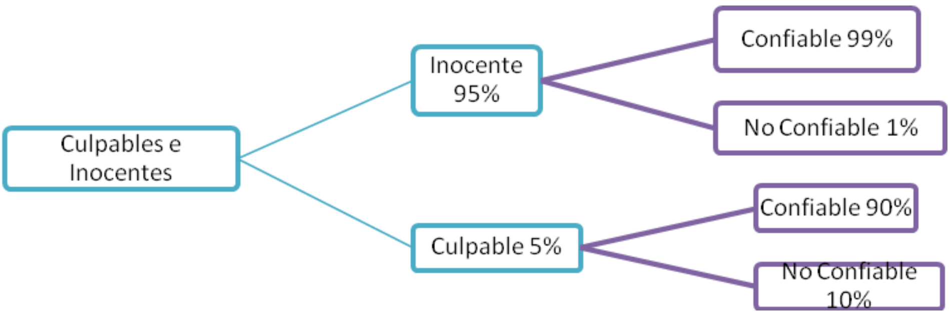


Ilustración 5.11 Gráfico del caso suero de la verdad

Inocente

Confiable 99% = 0.95 x 0.99 = 0.9405

No Confiable 1% = 0.95 x 0.01 = 0.0095

Culpable 5%

Confiable 90% = 0.05 x 0.90 = 0.045

No Confiable 10% = 0.05 x 0.10 = 0.005

¿Cuál es la probabilidad de que sea inocente dado que él dijo que era culpable?

$$P(A/B) = P(A \cap B) / P(B)$$
$$P(A \cap B) = 0.0095 + 0.045 = 0.0545$$

$$\frac{P(A \cap B)}{P(B)} = \frac{0.0095}{0.0095 + 0.045} = \frac{0.0095}{0.0545} = 0.17$$

La probabilidad de que sea inocente dado que él dijo que era culpable es del 17%.

5.7 Ejercicios propuestos

1.- Enumere los elementos que contienen cada uno de los siguientes espacios muestrales:

a) la serie de números enteros entre 1 y 50 que son divisibles entre 8.

b) el conjunto S = {x / x2 + 4x – 5 = 0}

c) el conjunto de lanzar una moneda al aire hasta que aparecen una cruz o tres caras.

d) el conjunto S = (x / x es un continente)

e) el conjunto S = {x / 2x – 4 ≥ 0 y x < 1}

2.- ¿Cuáles de los siguientes eventos son iguales?

a) A = {1, 3}

b) B = {x / x es un número de un dado}

c) C = {x / x2 – 4x + 3 = 0}

d) D = {x / x es el número de caras cuando se lanzan seis monedas al aire}

3.- De un conjunto de estudiantes de estadística se eligen cuatro al azar y se clasifican como hombre o mujer.

Enumere los elementos del espacio muestral S1 usando la letra H para hombre y M para mujer. Puntualice un segundo espacio muestral S2 donde los elementos representen el número de mujeres seleccionadas.

4.- Si S = {0, 1, 2, 3, 4, 5, 6, 7, 8, 9} y A = {0, 2, 4, 6, 8}, B = {1, 3, 5, 7, 9}, C = {2, 3, 4, 5} y D = {1, 6, 7}, liste los
elementos de los conjuntos que corresponden a los siguientes eventos:

a) $A \cup C$; b) $A \cap B$; c) $\neg C$; d) $(\neg C \cap D) \cup B$; e) $(S \cap \neg C)$; f) $A \cap C \cap \neg D$.

5.- Suponga el espacio muestral S = {potasio, uranio, oxígeno, cinc, cobre, sodio, nitrógeno,} y los eventos: A = {cobre, cinc, sodio}; B = {sodio, potasio, nitrógeno} y C = {oxígeno}.

Liste los elementos químicos de los conjuntos que pertenecen a los siguientes eventos:

a) $\neg A$; b) $A \cup C$;c) $(A \cap \neg B) \cup \neg C$; d) $\neg B \cap \neg C$;

e)$A \cap B \cap C$; f) $(\neg A \cup \neg B) \cap (\neg A \cap C)$.

6.- Sean los siguientes eventos A, B y C pertenecientes al espacio muestral S. Realice diagramas de Venn para sombrear las zonas que representan los siguientes eventos:

a) $(A \cap B)$; b) $(A \cup B)$; c) $(A \cap C) \cup B$.

6.- Una nueva medicina para aliviar la hipertensión se puede adquirir en 5 diferentes laboratorios y en forma de comprimidos, líquido o cápsulas, todas las presentaciones tienen dos indicadores de concentración, normal o alta.

¿De cuántas maneras diferentes un galeno puede recetar la medicina a un paciente que sufre esta enfermedad?

7.- ¿De cuántas maneras distintas puede responder un estudiante que desea rendir una prueba de falso-verdadero que consta de 9 preguntas?

8.- De cuántas formas se pueden cubrir las 5 posiciones iniciales en un equipo de fútbol con 8 jugadores que pueden jugar cualquiera de las posiciones?

9.- ¿Cuántas permutaciones diferentes se logran hacer con la literatura de la palabra COLUMNA? ¿Cuántas de estas permutaciones inician con la letra M?

10.- ¿Cuántas permutaciones diferentes se pueden lograr con las letras de la palabra INFINITO?

11.- Las probabilidades de que una maquinaria procesadora de alimentos cometa es 0, 1, 2, 3, 4 o más errores al producir un producto son 0.19, 0.34, -0.25, 0.43 y 0.29, respectivamente.

12.- La probabilidad de que una industria electrónica Japonesa se ubique en Quito, Ecuador, es 0.7, la probabilidad de que se ubique en Guayaquil, es 0.4 y la probabilidad de que se sitúe en Quito o Guayaquil, o en ambas ciudades, es 0.8.

Cuál es la probabilidad de que la industria se ubique…

a) en ambas ciudades?
b) en ninguna de esas ciudades?

13.- En cierta región del país se conoce por experiencia que la probabilidad de seleccionar un adulto mayor de 60 años de edad con cáncer es 0.05. Si la probabilidad de que un médico diagnostique de manera correcta que alguna persona con cáncer tiene la enfermedad es 0.78, y la posibilidad de que diagnostique de manera incorrecta que alguna persona sin cáncer tiene la enfermedad es 0.06. ¿Cuál es la probabilidad de que a un adulto mayor de 60 años se le determine cáncer?

14.- La posibilidad de que un paciente se recupere de una delicada intervención de corazón es 0.8. ¿Cuál es la probabilidad de que… a) puntualmente 2 de los siguientes 3 pacientes a los que se realiza esta intervención sobrevivan? b) Los siguientes 3 pacientes que tengan esta intervención sobrevivan?

15.- Si se tienen 5 peras verdes, 4 marrones y 6 amarillas, ¿cuántas opciones de 9 peras se pueden hacer si se deben seleccionar 3 de cada color?

16.- Un cargamento de 12 computadores contiene 3 defectuosos. ¿De cuántas formas puede una Sala de Cómputo comprar 5 de estos dispositivos y recibir al menos 2 defectuosos?

17.- El control de calidad a menudo se usa la ciencia estadística para determinar si un proceso está fuera de control. Suponga que el proceso, de hecho, está fuera de control y que 20% de los artículos producidos tiene imperfecto. a) Si tres productos salen en serie de la línea de producción, ¿cuál es la probabilidad de que los tres estén imperfectos? b) Si salen cuatro productos en serie, ¿cuál es la probabilidad de que tres estén imperfectos?

18.- En una conferencia se reúnen 250 médicos, de los cuales 115 son cardiólogos; 65, neurólogos y 70 otorrinolaringólogos. De estos médicos, el 75% de los cardiólogos, el 60% de los neurólogos y el 65% de los otorrinolaringólogos están a favor de utilizar una nueva anestesia para las cirugías. Si escogemos un médico al azar, y está a favor de aplicar la anestesia, ¿cuál es la probabilidad de que sea neurólogo?

19.- Un experimento en un invernadero inteligente suministrado de sistemas de energía solar y eólica, se sabe que la energía procesada cada día procede de paneles solares con una posibilidad de 0.4, de molinos eólicos con 0.26 y de ambos tipos de instalaciones con el 0,12. Escogiendo un día al azar, calcule la probabilidad de que tal energía procesada sea suministrada al invernadero:
a) por alguna de las dos instalaciones.

b) solo por una de las dos.

5.8 Ecuaciones introducidas en el capítulo 5
La multiplicación
Se da cuando los sucesos o eventos, ocurren el uno a continuación del otro, es como un suceso compuesto.

[5.1] $$\mathbf{n_1 * n_2 * n_3 * \ldots * n_x}$$

Factorial de un número
La factorial de un número natural ***n*** mayor que uno (1) es igual al producto de los ***n*** principales números naturales; el símbolo característico es "!". Así:

[5.2] $$\boldsymbol{n}! = \boldsymbol{n}!\ \forall\ (\boldsymbol{n} \in \boldsymbol{R})$$

Permutación
La permutación es el arreglo de todos los objetos del conjunto o parte de estos, a la permutación le importa el orden de los elementos involucrados.

[5.3] $$\mathbf{{}_nP_r} = \boldsymbol{n}!$$

Permutación sin repetición
Si en el arreglo de los elementos de la permutación corresponda a una parte de ellos

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

Permutación con repetición

Se llama permutación con repetición, al número de permutaciones de distintos objetos en la que posee elementos repetidos de primer orden, de un segundo orden, hasta de enésimo orden y que este cumpla el orden de los eventos.

[5.5] $${}_{n}P_{r1,\, r2,\, r3} = \frac{n!}{r_1! + r_2! + \cdots + r_k!}$$

Combinaciones

Para las combinaciones son importantes las formas posibles de seleccionar r objetos de un total de n, sin importar el orden. El número de combinaciones de n objetos distintos, tomando r a la vez.

[5.6] $${}_{n}C_{r} = \frac{n!}{r!(n-r)!}$$

Probabilidad

Si un experimento consigue obtener como resultado cualquiera de N diferentes resultados igualmente probables, y si exactamente en n de estos resultados comprenden al evento A.

[5.7] $$P(A) = \frac{n}{N}$$

Regla especial de la adición

Para emplear la regla especial de la adición, los eventos obligatoriamente tienen que ser mutuamente excluyentes. Acordarse que mutuamente excluyentes significa que cuando un evento sucede, ninguno de los demás eventos puede pasar al mismo tiempo.

[5.8] $$\mathbf{P\ (A\ U\ B) = P\ (A) + P\ (B)}$$

Regla general de la adición

Los resultados de un experimento pueden no ser mutuamente excluyentes.

[5.9] $$\mathbf{P\ (A \cup B) = P\ (A) + P\ (B) - P\ (A \cap B)}$$

Regla del complemento

Se pide para establecer la probabilidad de que un evento ocurra restando de 1 la probabilidad de un evento que no ha ocurrido.

[5.10] $$P(A) = 1 - P(\neg A) \quad \equiv \quad P(A) + P(\neg A) = 1$$

(Es equivalente)

Regla especial de la multiplicación

La regla especial de la multiplicación pide que dos eventos, A y B, sean independientes y lo son si el hecho de que uno suceda no altera la probabilidad de que el otro acontezca.

[5.11] $$\mathbf{P\ (A \cap B) = P\ (A) * P\ (B)}$$

Regla general de la multiplicación

Si dos eventos no son independientes, se dice que son dependientes, es decir sirve para determinar la probabilidad conjunta de dos eventos cuando éstos no son independientes.

[5.12] $$P\ (A \cap B) = P\ (A) * P\ (B/A) \quad \equiv \quad P\ (B/A) = \frac{P\ (A \cap B)}{P\ (A)}$$

(Es equivalente)

<u>Teorema de Bayes</u>

El teorema de Bayes es una situación probabilística en la es posible conocer las probabilidades de que acontezcan una serie de sucesos.

[5.13] $$\mathbf{P\ (A\ /B) = \frac{P\ (\ A\ ) * P\ (B/A)}{\sum [P(\ A\ ) * P(B/A)]}}$$

Unidad 6

VARIABLES ALEATORIAS Y DISTRIBUCIÓN CONJUNTA

"Debemos admitir con humildad que, mientras el número es puramente un producto de nuestra mente, el espacio tiene una realidad fuera de nuestra mente, de modo que no podemos prescribir completamente sus propiedades a priori."

-CARL FRIEDRICH GAUSS-

Matemático, astrónomo y físico alemán considerado "el príncipe de las matemáticas", es uno de los matemáticos más importantes de la historia.

CONTENIDO DE LA UNIDAD

INTRODUCCIÓN

En la teoría de la probabilidad (revisada en el capítulo anterior) se ilustra de manera sencilla la relación entre intuiciones subjetivas y estructuras objetivas de probabilidad. La paradoja mencionada muestra con frecuencia una contradicción entre el modelo matemático y la representación intuitiva de la situación. Por eso, un principio de aprendizaje para la unidad es identificar y discernir los modelos probabilísticos, correspondiente a cada caso con respecto a su variable aleatoria. Lo importante es el aprendizaje simultáneo en un nivel empírico-experimental y en un nivel teórico del modelo.

OBJETIVOS DE APRENDIZAJE

Al finalizar el capítulo, será capaz de:

- Relacionar probabilidad y variable aleatoria.
- Reconocer y distinguir entre variables continuas y discretas.
- Expresar los conceptos de Dependencia e independencia de variables aleatorias.
- Conocer las características principales de cada modelo de distribución.
- Identificar las características y uso de las distribuciones de probabilidad continua.

La gran importancia de esta unidad obedece a que al momento de realizar un experimento con un espacio muestral, siempre habrá situaciones en las que las variables aleatorias den como resultado parámetros sujetos a eventos de dependencia e independencia, esto permitirán posteriormente al estudiante profundizar en esta disciplina. Adicionalmente los temas de esta asignatura habilitarán al estudiante a desarrollar nuevo conocimiento en este contenido.

6.1 Definición de variables aleatorias

La correcta definición de una variable aleatoria es una función que relaciona un número real con cada elemento del espacio muestral. Utilizaremos una letra mayúscula, digamos X, para indicar una variable aleatoria y su oportuna letra minúscula, x en este caso, para uno de sus valores

.

6.1.1 Características de una distribución de probabilidad

Dada la situación:

1. La posibilidad de un resultado en específico se encuentra entre 0 y 1, inclusive.
2. Los posibles resultados son partes de eventos mutuamente excluyentes.
3. La adición de las probabilidades de los varios eventos es igual a 1.

Ejemplo 6.1: En un bingo, la urna del sorteo de los premios contiene 4 esferas rojas y 3 negras se sacan 2 esferas de manera sucesiva, sin reemplazo. Las posibles consecuencias y los valores y de la variable aleatoria Y, donde Y es el número de esferas rojas, son:

Tabla 6.1

Probabilidad del caso

ESPACIO MUESTRAL	y
RR	2
RN	1
NR	1
NN	0

6.1.2 Clasificación de las variables aleatorias

Se dice que una distribución de probabilidades muestra los resultados esperados al realizar un experimento, junto con la probabilidad en cada uno de estos resultados. Es decir, nos referimos a los valores posibles de una variable con sus respectivas probabilidades.

Estas distribuciones de probabilidad pueden corresponder a variables aleatorias discretas o contínuas

Variables aleatorias discretas: Se llaman aleatorias discretas a un espacio muestral posee un número finito de posibilidades, o una serie interminable con tantos elementos como números enteros existen.

Entre las distribuciones discretas existen los modelos más utilizados como: ***Uniforme, Bernoulli, Binomial, Binomial Negativa, Multinomial, Geométrica, Hipergeométrica, y Poisson.***

Variables aleatorias contínuas: Se le denomina espacio muestral continuo, si el espacio muestral posee un número infinito de posibilidades, igual al número de puntos en un segmento de recta.

Entre las distribuciones contínuas existen los modelos más utilizados, que son: ***Uniforme, Gamma, Exponencial, Normal, T Student y Chi Cuadrado.***

6.2 Aleatorias discretas

Una variable aleatoria discreta brinda cada uno de sus objetos con cierto grado de probabilidad. El conjunto de pares ordenados ***(x, f (x))*** es una función de probabilidad, una función de probabilidad o una distribución de probabilidad de la variable aleatoria discreta X si, para cada resultado x posible.

- $F(x) \geq 0$
- $\sum f(x) = 1$
- $P(X = x) = F(x)$

6.2.1 Distribución uniforme

La distribución uniforme tiene exactamente la misma probabilidad para todos sus elementos, es en la cual la variable aleatoria asume cada uno de sus valores con idéntica probabilidad.

Para una variable aleatoria x con una distribución uniforme discreta, su distribución de probabilidad está dada por:

[6.1] $$F(x) = \frac{1}{x}\text{; Donde } x = x_1, x_2, x_3 \ldots\ldots\ldots x_n$$

Cálculo de la media y la varianza de una distribución uniforme:

[6.2] $$\text{Media: } \mu = E(x) = \frac{\sum_{i=1}^{k} x_i\, f(x)}{k}$$

[6.3] **Varianza:** $\sigma^2 = E(x)^2 - [E(x)]^2$

Características de la distribución uniforme

- La variable aleatoria adquiere una probabilidad idéntica para cada uno de sus valores.
- Mediante los parámetros, se otorga la inversa del (de los) valor(es) que puede tomar la variable aleatoria.
- La media siempre coincide con cualquiera de los valores observados en el análisis.
- La varianza depende de los valores que pueda tomar la media.

Ejemplo 6.2: Hallar el valor esperado y la varianza de la variable aleatoria x que representa el número de puntos al realzar un lanzamiento de un dado, donde x = 1, 2, 3, 4, 5, 6.

x	1	2	3	4	5	6
$f(x)$	1/6	1/6	1/6	1/6	1/6	1/6

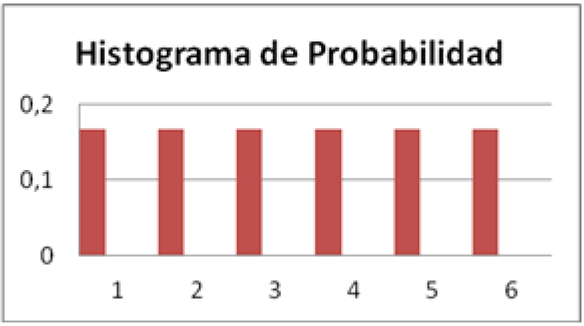


Ilustración 6.1: Distribución uniforme discreta

Media: $\mu = E(x) = \frac{\sum_{i=1}^{k} xi\, f(x)}{k} = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6}$

$$E(x)^2 = \sum_{i=1}^{k} xi^2\, f(x) = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6}$$

Varianza: $\sigma^2 = E(x)^2 - [E(x)]^2 = \frac{91}{6} - \left[\frac{21}{6}\right]^2 = 2{,}91$

Desviación Estándar: $\sigma = \sqrt{2.91} = 1.70$

Ejemplo 6.3: Hallar la variable aleatoria x posee la siguiente distribución de probabilidad. Encontrar desviación estándar.

x	-2	3	5
F(x)	0,3	0,2	0,5

$$\mu = E(x) = (-2)(0{,}3) + (3)(0{,}2) + (5)(0{,}5) = 2{,}5$$
$$E(x)^2 = 4(0{,}3) + 9(0{,}2) + 25(0{,}5) = 15{,}5$$
$$\sigma^2 = E(x)^2 - [E(x)]^2 = 15{,}5 - (2{,}5)^2 = 9{,}25$$
$$\sigma = \sqrt{9{,}25} = 3{,}04$$

6.2.2 Distribución de Bernoulli

La distribución de Bernoulli es un caso especial de la distribución binomial, donde ***n*** es el número de experimentos, ***x*** número de existo, ***p*** probabilidad de éxito y ***(1 − p)*** probabilidad de fracaso ***(q)***.

[6.4] $$b\,(x) = p * q^{\,1-p}$$

En técnicas precisas el proceso de Bernoulli se caracteriza por lo siguiente:

- El experimento consta de ensayos repetidos.
- Cada ensayo produce un resultado que se puede clasificar como éxito o fracaso.
- La probabilidad de un éxito, que se denota con p, permanece constante de un ensayo a otro.

- Los ensayos repetidos son independientes.

Ejemplo 6.4: Determine mediante el proceso de Bernoulli la probabilidad de obtener cinco caras al momento de lanzar una moneda doce veces.

X = 5; n = 12; p= 0,5; 1 - q = 0,5

$$P(x=5) = \binom{12}{5}(0{,}5)^5(0{,}5)^7 = 0{,}19 \rightarrow 19\%$$

Al momento de realizar los doce lanzamientos de la moneda, existe una probabilidad del 19% que salgan cinco veces cara.

6.2.3 Distribución binomial

La distribución de probabilidad binomial y de Bernoulli es una variable aleatoria discreta que se presenta con mucha frecuencia. Consiste en repetir un experimento n veces, todos los eventos son independientes entre sí, las probabilidades de éxito y fracaso permanecen constantes, los resultados del experimento o los resultados del evento solo pueden ser éxito o fracaso, esto es un rasgo principal de una distribución binomial, consiste en que solo hay dos posibles resultados en determinado intento de un experimento.

Una experimentación de variables discretas Binomial alcanza obtener como resultado un éxito con probabilidad **p** y un fracaso con probabilidad **q = 1 – p**. Por tal razón, la distribución de probabilidad de la variable aleatoria binomial **x**, es el número de éxitos en **n** ensayos independientes.

[6.5] $$B\,(x; n, p) = \binom{n}{x} p^n q^{n-x},\ x = 0, 1, 2 \ldots n.$$

La media y la varianza de la distribución binomial b (x; n, p) son:

[6.6] **Media:** $\mu = n*p$

[6.7] **Varianza:** $\sigma^2 = n*p*q$

Características de la distribución binomial

- Los experimentos tienen dos tipos de resultados, uno que es lo que se espera que ocurra o lo contrario, denominados "éxito" y "fracaso".

- Las probabilidades de estas deducciones son constantes, es decir no cambian.
- Cada ensayo o repetición, se diferencia el uno del otro porque son independientes entre ellos.
- El total de ensayos es una constante n.

Ejemplo 6.5: La probabilidad de que alguna clase de componente electrónico se averíe en una prueba de choque automovilístico es de $\frac{3}{4}$. Calcule la probabilidad de que se dañen exactamente 2 de los siguientes 4 componentes que se prueben. Si suponemos que las pruebas son independientes y p = $\frac{3}{4}$ para cada una de las 4 pruebas, obtenemos:

x = 2 (componentes dañados)
n = 4 (componentes totales)
p = $\frac{3}{4}$ (probabilidad éxito)
q = $\frac{1}{4}$ (fracaso)

$$B\left(2;\ 4, \frac{3}{4}\right) = \binom{4}{2}\left(\frac{3}{4}\right)^{4}\left(\frac{1}{4}\right)^{4-2} = 6 * 0{,}5625 * 0{,}0625 = 0.2109 \rightarrow 21{,}09\%$$

La probabilidad que se dañen 2 componentes electrónicos es del 21,09% de posibilidad.

Ejemplo 6.6: Se supone que hay contaminaciones en 30% del total de pozos de agua potable de cierta comunidad rural de Daule. Para obtener información sobre la verdadera dimensión del problema se establece que debe realizarse algún tipo de prueba. Como es muy costoso probar todos los pozos del área, se eligen 10 al azar para realizar la prueba.

a) Si se manipula la distribución binomial ¿cuál es la probabilidad de que debidamente 3 pozos tengan impurezas, considerando que la suposición es correcta?

b) ¿Cuál es la probabilidad de que más de 3 pozos tengan impurezas?

A) $b(3; 10{,}0{,}3) = \sum_{x=0}^{3} b(x; 10{,}0{,}3) - \sum_{x=0}^{2} b(x; 10{,}0{,}3)$

$$B(3; 10, 0{,}3) = 0{,}6496 - 0{,}3828 = 0.2668$$

B) P (X > 3) = 1 – 0,6496 = 0,3504

Ejemplo 6.7: La probabilidad de que un enfermo se recobre de un raro padecimiento sanguíneo es de 0.4. Si se sabe que 15 personas contrajeron la enfermedad, ¿cuál es la probabilidad de que a) sobrevivan al menos 10, b) sobrevivan de 3 a 8?; c) Calcule la media, la varianza y desviación estándar de la variable aleatoria binomial.

a) $P\ (x \geq 10) = 1 - P\ (x < 10)$

b) $P\ (x \geq 10) = 1 - \sum_{x=0}^{9} B(x; 15{,}0.4) = 1 - 0{,}9662 = 0{,}0338$

c) $P\ (3 \leq x \leq 8)$

$$\sum_{x=3}^{8} b(x; 15{,}0.4) = \sum_{x=0}^{8} b(x; 15{,}0.4) - \sum_{x=0}^{2} b(x; 15{,}0.4) = 0{,}9050 - 0{,}0271$$
$$= 0{,}8779$$

d) $\mu = 15*(0{,}4) = 6$ $\sigma^2 = (15)(0{,}4)(0{,}6) = 3{,}6$ $\sigma = \sqrt{3{,}6} = 1{,}90$

6.2.4 Distribución binomial negativa

Si pruebas independientes repetidas pueden dar como consecuencia un éxito con probabilidad p y un fracaso con probabilidad q = 1 – p, en aquella distribución de probabilidad de la variable aleatoria x, el número de pruebas en el que ocurre el k-ésimo éxito, es:

[6.8] $$B * (x;\ k, p) = \binom{x-1}{k-1} p^k * q^{x-k};\ k = k, k+1, k+2 \ldots$$

El número x de pruebas necesarias para formar k éxitos en un experimento binomial negativo se nombra variable aleatoria binomial negativa y su distribución de probabilidad se llama distribución binomial negativa.

[6.9] **Esperanza:** $\mu = \frac{(k-q)}{p}$

[6.10] **Varianza:** $\mu = \frac{(k-q)}{p^2}$

Sus probabilidades obedecen del número de éxitos deseados y de la probabilidad de un éxito en un experimento dado, indicaremos ambas probabilidades con el símbolo ***b*(x; k, p).*** Para alcanzar la fórmula general para ***b*(x; k, p)***, suponga que la probabilidad de un éxito en el ***x***-ésimo experimento precedido por $\boldsymbol{k-1}$ éxitos y $\boldsymbol{x-k}$ fracasos en un orden detallado. Como los experimentos son independientes podemos multiplicar todas las probabilidades que corresponden a cada resultado deseado, la probabilidad de que acontezca un éxito es $\boldsymbol{p}$ y la probabilidad de que suceda un fracaso es $\boldsymbol{q = 1 - p}$.

Características de la distribución binomial negativa

- Cada ensayo produce un resultado posible mutuamente excluyente que se puede clasificar como éxito o fracaso.
- La probabilidad de conseguir un resultado en cada una de las pruebas es de éxito o fracaso.
- Las probabilidades de éxito o fracaso son constantes.
- La distribución refleja de un número indefinido de pruebas independientes o separables.

Ejemplo 6.8: En la serie de campeonato europeo de la Champions se decreta que, el equipo que gane 4 de 7 juegos será el ganador. Suponga que los equipos Barcelona (A) y Real Madrid (B) se enfrentan en los juegos de campeonato y que el equipo A tiene una probabilidad de 0.55 de ganarle al equipo B.

a) ¿Cuál es la probabilidad que el equipo A domine la serie en 6 juegos?
b) ¿Cuál es la probabilidad que el equipo A gane la serie?
c) Si A y B equipos se afrontaran en la eliminatoria de la serie regional y el victorioso fuera el que ganara 3 de 5 juegos, ¿cuál es la probabilidad de que el equipo A gane la serie?

x = 6 (juegos dominados)
k = 4 (gane el equipo)
p = 0,55 (probabilidad de ganar)

a) *b* (6; 4, 0.55)* = $\binom{5}{3}(0{,}55)^4(1-0{,}55)^{6-4} = 0{,}1853$

b) P (equipo A gana la serie de campeonato)

P(A gane serie) = b* (4; 4, 0.55) +b* (5; 4, 0.55) +b* (6; 4, 0.55) +b* (7; 4, 0.55)

P(A gane)= 0.0915 + 0.1647 + 0.1853 + 0.1668 = 0.6083.

c) P (el equipo A gana la eliminatoria)

P(A gane eliminatoria) =b* (3; 3, 0.55) +b* (4; 3, 0.55) +b* (5; 3, 0.55)

P(A gane eliminatoria) = 0.1664 + 0.2246 + 0.2021 = 0.5931.

6.2.5 Distribución multinomial

La distribución del experimento binomial se transforma en el caso de en un experimento multinomial si cada prueba posee más de dos resultados posibles.

Si un ensayo dado que consigue producir los **k** resultados $\mathbf{E_1, E_2, \ldots, E_k}$ con probabilidades $\mathbf{p_1, p_2, \ldots, p_k}$, entonces la distribución de probabilidad de las variables aleatorias $\mathbf{X_1, X_2, \ldots, X_k}$, que constituye el número de ocurrencias para $\mathbf{E_1, E_2, \ldots, E_k}$ en ***n*** ensayos independientes, es:

[6.11] $$f(\boldsymbol{x_1, x_2 \ldots . x_k; p_1, p_2 \ldots p_k}) = \binom{\boldsymbol{n}}{\boldsymbol{x_1, x_2 \ldots . x_k}} \boldsymbol{p_1}^{\boldsymbol{x1}} * \boldsymbol{p_2}^{\boldsymbol{x2}} \ldots . \boldsymbol{p_k}^{\boldsymbol{xk}}$$

$$\sum_{x=1}^{k} \boldsymbol{x_i} = \boldsymbol{n}\ \boldsymbol{y} \sum_{x=1}^{k} \boldsymbol{p_i} = \mathbf{1}$$

Características de la distribución multinomial

- Cuando se lleva a cabo un experimento, este puede concluir con más de un resultado.
- Las probabilidades relacionadas a cada experimento poseen resultados constantes.
- La independencia en cada uno de los ensayos.
- El ensayo se puede repetir manera ilimitada.

Ejemplo 6.9: La complicación de los arribos y los despegues de los aviones en un aeropuerto es tal que a menudo se maneja la simulación por computadora para formar las condiciones ideales. Para un aeropuerto específico que posee tres pistas se conoce que, en el escenario ideal, las probabilidades de que las pistas individuales sean manejadas por un avión comercial que llega aleatoriamente son las siguientes:

- Pista 1: p1 $= \frac{2}{9}$
- Pista 2: p2 $= \frac{1}{6}$
- Pista 3: p3 $= \frac{11}{18}$

¿Cuál es la probabilidad de que 6 aviones que llegan al azar se distribuyan de la siguiente manera?

- Pista 1: 2 aviones.
- Pista 2: 1 avión.
- Pista 3: 3 aviones.

x = 2, 1, 3 (número de aviones).
$\mathbf{p} = \frac{2}{9}, \frac{1}{6}, \frac{11}{18}$ (probabilidad de las pistas).
n = 6 (número de aviones).

$$f\left(2, 1, 3; \frac{2}{9}, \frac{1}{6}, \frac{11}{18}, 6\right) = \binom{6}{2, 1, 3}\left(\frac{2}{9}\right)^2\left(\frac{1}{6}\right)^1\left(\frac{11}{18}\right)^3$$

$$f\left(2, 1, 3; \frac{2}{9}, \frac{1}{6}, \frac{11}{18}, 6\right) = \left(\frac{6!}{2! * 1! * 3!}\right)\left(\frac{2}{9}\right)^2\left(\frac{1}{6}\right)^1\left(\frac{11}{18}\right)^3 = 0{,}1127$$

6.2.6 Distribución geométrica

Si pruebas independientes frecuentadas pueden obtener como consecuencia un éxito con probabilidad p y un fracaso con probabilidad $q = 1 - p$, entonces la distribución de probabilidad de la variable aleatoria X, el número de la prueba en el que ocurre el primer éxito, es:

[6.12] $$G(x; p) = p * q^{x-1}, \; x = 1, 2, 3, \ldots n$$

Muy a menudo, en aplicaciones que tienen que ver con la distribución geométrica, la media y la varianza son importantes, la media y la varianza de una variable aleatoria que sigue la distribución geométrica son:

[6.13] **Media:** $\mu = \frac{1}{p}$

[6.14] **Varianza:** $\sigma^2 = \frac{1-p}{p^2}$

Características de la distribución geométrica

- Radica en pruebas a un número fijo, de n de pruebas.
- Cada ensayo produce un resultado que se puede clasificar como éxito o fracaso.
- Siempre hay una probabilidad de éxito valor p y de fracaso que es igual a q = 1-p.
- Cada prueba es independiente.

Ejemplo 6.10: Se sabe al realizar proceso de fabricación de una planta metalúrgica uno de cada 100 artículos, en promedio, resulta defectuoso. ¿Cuál es la probabilidad de que el quinto artículo que se examina, en un segmento de 100, sea el primer defectuoso que se encuentra?

x = 5 (defectuoso)
p = 0,01 (probabilidad que sea defectuoso)
q = 1-0,01= 0,99

$$G\ (5; 0{,}01) = (0{,}01)*(0{,}99)^4 = 0{,}0096.$$

La probabilidad de que el quinto artículo sea defectuoso de un grupo de 100, es del 0,96%.

Ejemplo 6.11: Un conmutador telefónico está en su límite de capacidad, por lo que los beneficiarios tienen dificultad para realizar sus llamadas. Sería atractivo saber cuántos intentos serían necesarios para obtener un enlace telefónico. Suponga que la probabilidad de alcanzar un enlace durante un momento agitado es p = 0.05. Si nos interesa conocer la probabilidad de que se necesiten 5 intentos para vincular con éxito una llamada. Calcule la media y la varianza de la variable aleatoria.

x = 5 y p = 0.05

P (X = x) = G (5; 0.05) = (0.05)(0.95)4 = 0.041.*

$$\mu = \frac{1}{0{,}05} = 20 \; \sigma^2 = \frac{1 - 0{,}05}{0{,}05^2} = 380$$

6.2.7 Distribución hipergeométrica

Los tipos de aplicaciones de la distribución hipergeométrica son muy similares a los de la distribución binomial. Nos interesa el cálculo de probabilidades para el número de observaciones que caen en una categoría específica.

La distribución de probabilidad de la variable aleatoria hipergeométrica **X**, el número de éxitos en una muestra aleatoria de tamaño n que se selecciona de **N** artículos, en los que **k** se denomina éxito y **N – k** fracaso.

[6.15] $$h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \; ; \max\left(0, n - (N-k)\right) \leq x \leq \min(n, k)$$

Cálculo de la media y la varianza de una distribución hipergeométrica:

[6.16] **Media:** $\mu = \frac{nk}{N}$

[6.17] **Varianza:** $\sigma^2 = \frac{N-n}{N-1} * n * \frac{k}{N} * (1 - \frac{k}{N})$

Por tanto, si se escoge una muestra de una población finita sin reemplazos y si el tamaño de la muestra **n** es mayor que **5%** del tamaño de la población, se aplica la distribución hipergeométrica para determinar la probabilidad de un número especifico de éxitos o fracasos. Esto resulta fundamentalmente apropiado cuando la dimensión de la población es pequeña.

Características de la distribución hipergeométrica

- Cada ensayo produce un resultado que se puede clasificar como éxito o fracaso.
- La variable aleatoria es el dígito de éxitos de un número fijo de pruebas.
- Las pruebas no son independientes.
- Los muestreos se ejecutan con una muestra finita sin reemplazos. Por tanto, la probabilidad de éxito cambia en cada prueba.

Ejemplo 6.12: Varios lotes con 40 dispositivos cada uno que contengan 3 o más imperfecciones se consideran inadmisibles. El procedimiento para obtener muestras del lote consiste en seleccionar 5 dispositivos al azar y rechazar el lote si se encuentra un componente imperfecto. ¿Cuál es la probabilidad de, que en la muestra, se encuentre exactamente un dispositivo imperfecto, si en todo el lote hay 3 imperfectos? Calcule la media y la varianza de la variable aleatoria.
n = 5 (dispositivos seleccionados)
N = 40 (lote de dispositivos)
k = 3 (dispositivos rechazados)
x = 1, encontremos

$$h(1; 40,5,3) = \frac{\binom{3}{1}\binom{40-3}{5-1}}{\binom{40}{5}} = 0.3011 \ \rightarrow 30.11\%$$

Este plan no es adecuado porque solo 30% de los intervalos detecta un lote malo con 3 dispositivos imperfectos.

$$\mu = \frac{(5)(3)}{40} = \frac{3}{8} = 0{,}375 \ ; \ \sigma^2 = \left(\frac{40-5}{40-1}\right) * (5) * \left(\frac{3}{40}\right) * \left(1 - \frac{3}{40}\right) = 0{,}3113$$

Ejemplo 6.13: Se usa aleatoriamente a 10 individuos para un estudio de caso médico. El grupo posee 3 individuos con sangre tipo O, 4 con sangre tipo A y 3 con sangre tipo B. ¿Cuál es la probabilidad de que una muestra aleatoria de 5 contenga 1 persona con sangre tipo O, 2 personas con tipo A y 2 personas con tipo B?

$$h(1,2,2; 3,4,3;\ 10,5) = \frac{\binom{3}{1}\binom{4}{2}\binom{3}{2}}{\binom{10}{5}} = \frac{3}{14} = 0{,}214 \ \rightarrow 21{,}42\%$$

6.2.8 Distribución de Poisson

La distribución de Poisson pertenece a las distribuciones de probabilidad de una variable de tipo discreta, esta brinda una serie de posibles ocurrencias dentro de un suceso por un número indeterminado de veces en un intervalo específico de medida calculable, que puede corresponder a: área, tiempo, distancia, volumen.

Características de la distribución de Poisson
La distribución de Poisson debe desempeñar las siguientes restricciones:

- La variable aleatoria X es las veces que el evento se repite durante un intervalo.
- Los sucesos deben ser aleatorias.
- Los eventos tienen que ser independientes entre sí.
- Los eventos deben estar idénticamente distribuidas adentro del intervalo que se emplea.
- La distribución de Poisson únicamente se ve afectada por el valor de la media μ, a la paradójica distribución binomial que se veía afectada por el tamaño de la muestra y la probabilidad.

Aplicaciones

- Defectos de nacimiento y mutaciones genéticas.
- Las enfermedades raras (como la leucemia, pero no el SIDA porque es infecciosa y por tanto no independiente).
- Accidentes automovilísticos.
- El flujo de tráfico y la distancia de seguridad.
- Número de errores de escritura en una página.
- Propagación de un animal en peligro de extinción en África.
- Fallos de una máquina en un mes.

La representación de la su función de probabilidad está proporcionada por la siguiente fórmula:

[6.18] $$f(\boldsymbol{x} = \boldsymbol{k}) = \left(\frac{\boldsymbol{e}^{-\mu}(\mu)^{\boldsymbol{x}}}{\boldsymbol{x}!}\right)$$

donde:

k: observación planteada.

e: logaritmo neperiano 2,71828.

x: número de ocurrencias.

μ: la medida de cantidad de veces de éxito que presenta un evento en un intervalo particular, es decir el promedio de ocurrencia de evento.

Ejemplo 6.14: Un hospital de la ciudad de Guayaquil analiza los nacimientos de bebés varones. Si se conoce que en una semana nacen una media de siete varones, determinar:

Probabilidad de que nazcan tres varones en una semana.

x= nacimientos de bebés varones.

μ = nacimiento de siete bebes varones a la semana.

$$P(x=3)=\left(\frac{e^{-7}(7)^{3}}{3!}\right)=0{,}052$$

Lo que significa que la probabilidad de que nazcan tres varones en una semana, será del 5,2%
Probabilidad de que nazcan menos de tres varones a la semana.

P(X<3) = P(x=0) + P(x=1) + P(x=2)
$\mathrm{P}(\mathrm{X}<3)=\left(\frac{e^{-7}(7)^{0}}{0!}\right)+\left(\frac{e^{-7}(7)^{1}}{1!}\right)+\left(\frac{e^{-7}(7)^{2}}{2!}\right)=0{,}029$

Lo que manifiesta que la probabilidad de que nazcan menos de tres bebés varones en una semana, será del 2,9%

Si dentro de las ocurrencias de un suceso μ se presentan intervalos, μ pasa a determinarse por **λ**t, donde **λ** representa la razón media por unidad y **t** el número de veces que ocurre, por lo cual la función de Poisson se puede también representar:

[6.18A] $$P(\boldsymbol{x};\ \boldsymbol{\lambda t})=\left(\frac{\boldsymbol{e^{-\lambda t}(\lambda t)^{x}}}{\boldsymbol{x}!}\right)$$

Ejemplo 6.15: En una fábrica de textiles, presenta un promedio μ de tres accidentes al mes, determine:

a) La probabilidad de que no ocurra ningún accidente al mes.

μ = 3 accidentes por mes

x = 0

$$P(\boldsymbol{x} = \mathbf{0}) = \left(\frac{\boldsymbol{e}^{-3}(\mathbf{3})^{\mathbf{0}}}{\mathbf{0}!}\right) = 0{,}04978$$

Existe la probabilidad del 4,98% de que no ocurra algún accidente.

a) De que ocurra como máximo dos accidentes al mes .

μ = 3 accidentes por mes

x=0, 1, 2.

$$P(x \leq 2) = P(x=0) + P(x=1) + P(x=2)$$

$$P(\boldsymbol{x}) = \left(\frac{\boldsymbol{e}^{-3}(\mathbf{3})^{\mathbf{0}}}{\mathbf{0}!}\right) + \left(\frac{\boldsymbol{e}^{-3}(\mathbf{3})^{\mathbf{1}}}{\mathbf{1}!}\right) + \left(\frac{\boldsymbol{e}^{-3}(\mathbf{3})^{\mathbf{2}}}{\mathbf{2}!}\right) = 0{,}3485$$

Existe la probabilidad del 34,85% de que ocurran dos accidente.

b) De que ocurra 30 accidentes en un año.

μ = 3 accidentes por mes

como el análisis ahora se mide por año y no por mes, la media μ corresponderá a:

$\mu = \lambda t$, por lo que resultará

3 accidente $= \lambda$ y 1 mes = t

μ = 3 accidentes * 12 meses = 36 accidentes en un año.

$$P(\boldsymbol{x};\, \boldsymbol{\lambda t}) = \left(\frac{\boldsymbol{e}^{-36}(\mathbf{36})^{\mathbf{30}}}{\mathbf{30}!}\right) = 0{,}052$$

c) De que ocurra 8 accidentes en un trimestres.

$\mu = \lambda t$; μ= 3 accidentes *3 meses = 9

x=8

$$P(\boldsymbol{x};\, \boldsymbol{\lambda t}) = \left(\frac{\boldsymbol{e}^{-9}(\mathbf{9})^{\mathbf{8}}}{\mathbf{8}!}\right) = 0{,}1317$$

Ejemplo 6.16: Durante un ensayo experimental de laboratorio el número promedio de átomos radiactivos que van a través de un contador en un milisegundo es 4. ¿Cuál es la probabilidad de que entren 6 átomos al contador en un milisegundo dado?
Donde:
x = 6 (átomos radioactivos)
$\boldsymbol{\lambda}$t = 4 (promedio)

$$P(6;\ 4) = \left(\frac{e^{-4}(4)^6}{6!}\right) = 0{,}1042$$

La probabilidad que ingresen 6 átomos es del 10,42%.

Ejemplo 6.17: En una fábrica los accidentes suceden con rara frecuencia. Se conoce que la probabilidad de un suceso accidental en cualquier día dado es de 0,005 y que los accidentes son independientes entre sí.

a) ¿Cuál es la probabilidad de que un día de cualquier tiempo definitivo de 400 días ocurra un accidente?

$$P(\boldsymbol{x} = \mathbf{1}) = \left(\frac{\boldsymbol{e}^{-\mathbf{2}}(\mathbf{2})^{\mathbf{1}}}{\mathbf{1}!}\right) = 0{,}271$$

b) ¿Cuál es la probabilidad que suceda un accidente a lo mucho en tres días de tal período?

$$P(\boldsymbol{x}) = \left(\frac{\boldsymbol{e}^{-\mathbf{2}}(\mathbf{2})^{\mathbf{0}}}{\mathbf{0}!}\right) + \left(\frac{\boldsymbol{e}^{-\mathbf{2}}(\mathbf{2})^{\mathbf{1}}}{\mathbf{1}!}\right) + \left(\frac{\boldsymbol{e}^{-\mathbf{2}}(\mathbf{2})^{\mathbf{2}}}{\mathbf{2}!}\right) = 0{,}857$$

6.3 Aleatorias contínuas

Estas distribuciones se establecen en variables aleatorias discretas, que sólo acogen valores notoriamente separados. En otras palabras, en el asunto de la distribución contínua, tal vez ansíe saber la participación de observaciones que se muestran adentro de cierto margen.

Es importante señalar que una variable aleatoria contínua posee un número infinito de valores mantenidos dentro de algún intervalo particular. Así, corresponde idealizar la probabilidad de que una variable posea un valor adentro de un intervalo específico, en vez de especular en la probabilidad de un valor específico, no importa que se incluyan o no el punto final o inicial en el intervalo.

6.3.1 Distribución uniforme contínua

La parte de distribución contínua más sencilla de la estadística es la distribución uniforme contínua. Tal es que en esta distribución se determina por una función de densidad que es "plana", por lo cual la probabilidad es de forma uniforme en un intervalo cerrado, digamos **[A, B]**, no obstante las concentraciones de la distribución uniforme contínua, no son tan abundantes como las de distintas distribuciones que se muestran en este capítulo.

La función de consistencia de la variable uniforme contínua X en el intervalo
[A, B] es

[6.19]
$$f(\boldsymbol{x}; \boldsymbol{A}, \boldsymbol{B}) = \begin{cases} \dfrac{1}{B-A}, & A \leq x \leq B \\ 0, & en\ otros\ casos \end{cases}$$

La función de consiste en una forma de un rectángulo con base $\boldsymbol{B - A}$ y altura constante $\frac{1}{B-A}$, como tal consecuencia, la distribución uniforme muchas veces se conoce como distribución rectangular, sin embargo, se debe notar que el intervalo continuamente no es cerrado: [A, B]; también puede ser (A, B)

.

La media y la varianza de la distribución contínua uniforme son:

[6.20] **Media:** $\mu = \frac{A+B}{2}$

[6.19] **Varianza:** $\boldsymbol{\sigma^2} = \frac{(B-A)^2}{12}$

Ejemplo 6.18: Imagine que el tiempo máximo que se puede reservar un salón de conferencias magna de alguna empresa motivadora son cuatro horas. Con mucha periodicidad tienen conferencias extensas y breves. Realmente, se puede opinar que la duración X de una conferencia tiene una distribución uniforme en el intervalo [0, 4].

a) ¿Cuál es la función de densidad?

b) ¿Cuál es la probabilidad que cualquier conferencia establecida perdure al menos 3 horas?

a) $f(x) = \begin{cases} \frac{1}{4}, & 0 \le x \le 4 \\ 0, & en\ otros\ casos \end{cases}$

b) $P[x \ \ge 3] = \int_3^4 \frac{1}{4} dx = \frac{1}{4}$

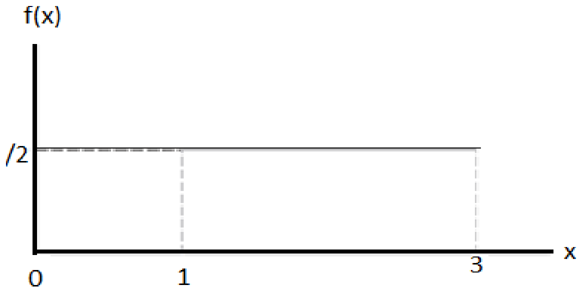


Ilustración 6.12 Distribución uniforme contínua

Ejemplo 6.19: Una compañía que suministra servicio de transporte de autobuses a las personas mientras se encuentran en un campus universitario. Un autobús arriba a la parada cada 30 minutos, entre las 6 de la mañana y las 11 de la noche toda la semana. Los estudiantes alcanzan a la parada en tiempos aleatorios. El tiempo que aguarda un estudiante tiene una distribución uniforme de 0 a 30 minutos.

a) Demuestre el área de la distribución uniforme.
b) ¿Cuánto tiempo esperará el autobús un estudiante? En otras palabras, ¿cuál es la media del tiempo de espera y la desviación estándar de los tiempos de espera?

a) $f(x) = \frac{1}{(30-0)} * (30 - 0) = 1$

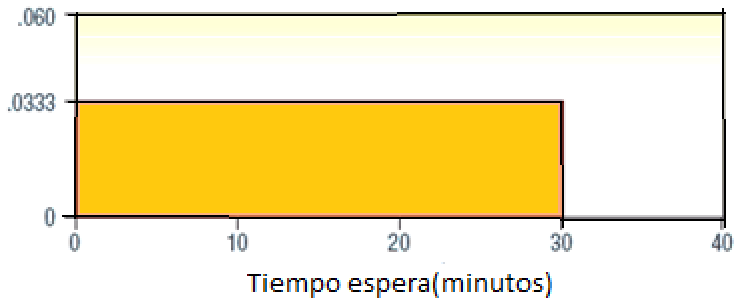


Ilustración 6.3 Servicio de transporte

b) $\mu = \frac{0+30}{2} = 15$ y $\sigma^2 = \frac{(30-0)^2}{12} = \sqrt{75} = 8.66$

La media de la distribución del tiempo es de 15 minutos y 8.66 minutos de la desviación estándar de la distribución.

$$f(x) = \begin{cases} \frac{c}{\sqrt{x}}, & 0 \leq x \leq 4 \\ 0, & en\ otros\ casos \end{cases}$$

Ejemplo 6.20: Dada la función f(x) donde el valor de la variable x sea un intervalo de 0 a 4;

a) Determinar el valor de c de manera que f(x) se considere en función de densidad de probabilidad.

b) Hallar P (x < 2) y P (X > 2).

a) $f(x) = \int_0^4 \frac{c}{\sqrt{x}} dx = 1$; condición para que la variable sea aleatoria.

$$c \int_0^4 \frac{1}{\sqrt{x}} dx = 1 \rightarrow 2c\sqrt{x}\Big|_0^4 = 1$$

$$4c = 1 \rightarrow c = 1/4$$

6.3.2 Distribución Gamma

La distribución Gamma se trata de una parte de distribuciones que provee un modelo adecuado para histogramas que presentan cierto tipo de asimetría. Antes de presentar a las variables aleatorias con distribución Gamma, es necesario recordar cómo se define la función Gamma o factorial, la cual cumple un rol importante en muchas ramas de la Matemática.

La variable Gamma contínua **X** posee una distribución gamma, con cuantificaciones α y β, si su función de densidad está dada por:

$$f(x;\propto,\beta) = \begin{cases} \frac{1}{\beta^{\propto}\,\Gamma(\propto)}\, x^{\propto-1} e^{-x/\beta}, & 0 > x \\ 0, & en\ otros\ casos \end{cases}$$

Donde $\alpha > 0$ y $\beta > 0$.

Características de la distribución Gamma

Las propiedades básicas de la distribución Gamma son:

- Si x1,...,xn son n variables aleatorias independientes distribuidas según una N(0,1).
- La nueva variable aleatoria Y= x21,..., x2n
- Sigue una distribución de la función gamma Γ(n)

Cuando el parámetro **P** es entero, a la distribución ***Γ(p,a)*** se le conoce con el nombre de distribución ***Erlang***

$$\Gamma(a)=\int_{0}^{\infty} x^{\propto-1}e^{-x}dx \; ; \; a > 0$$

Aplicación

La distribución gamma se suele utilizar en:

- Intervalos de tiempos entre dos fallos de máquinas, procesos, eléctricos, etc.
- Intervalos de tiempos entre dos llegadas de automóviles a un punto determinado.
- Tiempos de existencia de sistemas computacionales, mecánicos, etc.

Estadísticos:

[6.23] **Media:** $\frac{p}{a}$

[6.24] **Varianza:** $\frac{p}{a^2}$

Ejemplo 6.21: Un Call Center de servicios llegan 12 llamadas por minuto, siguiendo una distribución gamma. ¿Cuál es la probabilidad de que en menos de 1 minuto lleguen 8 llamadas?

$$P(x < 1) = \frac{1}{(\frac{1}{12}) * 2}\int_{0}^{1} u^{2-1}e^{-\frac{u}{1/12}}du$$

$$P(x < 1) = 0{,}9105$$

Existe un 91,05% de probabilidades de recibir 8 llamadas en un plazo de tiempo de menos de 1 minuto.

Ejemplo 6.22: Si se sabe que el tiempo de sobrevivencia de ratas expuestas a un determinado tóxico es una variable aleatoria que sigue una distribución Gamma (5, 10), ¿cuál es la probabilidad de que una rata no supere las 60 semanas de vida?

$$P(x < 60) = \frac{1}{10^5(5-1)!}\int_0^{60} x^{5-1}e^{-\frac{x}{10}}dx$$

$$P(x < 60) = 0{,}7149$$

6.3.3 Distribución de Exponencial

La distribución exponencial es un caso especial de la distribución gamma. La relación entre la gamma y la exponencial permite que la distribución gamma se utilice en tipos similares de problemas.

Frecuentemente se involucra la distribución exponencial en temas relacionados como al tiempo entre las llegadas, en las disposiciones de servicio, el tiempo de falla de los componentes y sistemas automatizados.

La variable aleatoria contínua X tiene una distribución exponencial, con parámetro β, si su función de densidad es dada por:

[6.25] $$f(x;\beta) = \begin{cases} \frac{1}{\beta}\, e^{-x/\beta}, & 0 > x \\ 0, & en\ otros\ casos \end{cases}$$

Donde $\beta > 0$

$$f(x) = \int_0^{x} \propto e^{-\propto x}dx = [-e^{-\propto x}] = 1 - e^{-\propto x}$$

La media y la variancia de la distribución exponencial son:

[6.26] **Media:** $\mu = \beta$

[6.27] **Varianza:** $\sigma^2 = \beta^2$

Características de la distribución exponencial

Las características básicas de la distribución Exponencial son:

- La esperanza de la desviación es σ.

- La media es el parámetro β.
- Si la variable x tiene como medida el tiempo, entonces pertenece a una distribución Exponencial.
- El parámetro de una distribución de Poisson λ (proceso de Poisson) con un número de sucesos por unidad de tiempo entre dos sucesos sucesivos sigue una distribución Exponencial donde el parámetro es α = 1/λ.

Ejemplo 6.23: Suponga que un sistema automatizado tiene cierto tipo de dispositivo cuyo tiempo de falla en años está dado por la variable aleatoria T, distribuida exponencialmente con tiempo promedio de falla β. Sí 5 de estos componentes se instalan en diferentes sistemas, ¿cuál es la probabilidad de que al menos 2 continúen funcionando después de 8 años?

$$p(x > 8) = \frac{1}{5}\int_{8}^{\infty} e^{-\frac{t}{5}}dt = e^{-\frac{8}{5}} = 0{,}2$$

n = 5

p = 0.20 = probabilidad de que un componente esté funcionando después de 8 años

q = 1 - p = 0,80 = probabilidad de que un componente no funcione después de 8 años

$$P(x \geq 2) = p(x = 2) + p\,(x = 3) + p(x = 4) + p\,(x = 5) = 1 - p\,(x = 0, 1)$$

$$P(x \geq 2) = 1 - [{}_5C_0 * (0{,}2)^0(0{,}8)^5 + {}_5C_1\,(0{,}2)^1(0{,}8)^4] = 1 - 0{,}7373 = 0{,}2627$$

Relación con la distribución de Poisson

Los estudios más importantes de la distribución exponencial son aquellos escenarios en donde se aplica el teorema de Poisson, es necesario recordar que un proceso de Poisson se maneja para calcular la probabilidad de un número específico de "eventos" en el transcurso de un período o espacio particular.

En muchos eventos, la cantidad de espacio o período es la variable aleatoria. Por ejemplo, una minera puede interesarse en el tiempo T entre arribos de sus vehículos en una intersección congestionada durante la hora de salida de trabajo en una cantera. Una experimento que representa el evento de Poisson.

La concordancia entre la distribución exponencial y la distribución de Poisson es bastante simple, la distribución de Poisson se amplió como una distribución de un solo parámetro**λ**, donde **λ** puede representarse como el número de eventos por unidad (tiempo). Suponiendo que ahora la variable aleatoria representada por el tiempo que se solicita para que ocurra el primer evento.

$$p(0, \lambda t) = \frac{e^{-\lambda t}(\lambda t)^0}{0!} = e^{-\lambda t}$$

Ejemplo 6.24: Sea Y una variable aleatoria contínua que tiene la siguiente función de densidad:

$$\mathrm{f}(y) = \begin{cases} c + \dfrac{3y}{8}, & 0 < y < 2 \\ 0, & en\ otros\ casos \end{cases}$$

Hallar:

a) El valor de c para que $\boldsymbol{f}$ **(y)** sea una función de densidad.

b) Obtener la función de distribución.

c) Calcular: **P (1 ≤ Y ≤ 1.5).**

d) Calcular: **P (Y > 1).**

a) Empleamos la expresión de función de densidad de probabilidad:

$$P(\propto \leq x \leq b) = \int_a^b \mathrm{f}(x)\, dx = 1$$

Para los intervalos expuestos en el enunciado del problema, en este caso:

$$P(0 \leq x \leq 2) = \int_0^2 \left(c + \frac{3y}{8}\right) dy$$

$$\int_0^2 cd\, dy + \int_0^2 \frac{3y}{8} dy = cy|_0^2 + \frac{3}{8} y^2|_0^2 = 2c + \frac{3}{4}$$

$$P(0 \leq x \leq 2) = 2c + \frac{3}{4} = 1;\ c = \frac{1}{8}$$

Posteriormente, obtenemos el valor de la constante **c** y la solución del problema, para que $\boldsymbol{f}$ **(y)** sea una función de probabilidad, el valor de **c** es de **1/8**. Por lo tanto, la densidad de la variable aleatoria continúa **Y** queda:

$$f(y) = \begin{cases} \frac{1}{8} + \frac{3y}{8}, & 0 < y < 2 \\ 0, & en\ otros\ casos \end{cases}$$

La representación gráfica de la función de densidad de probabilidad, ***f*** **(y)**, se muestra a continuación:

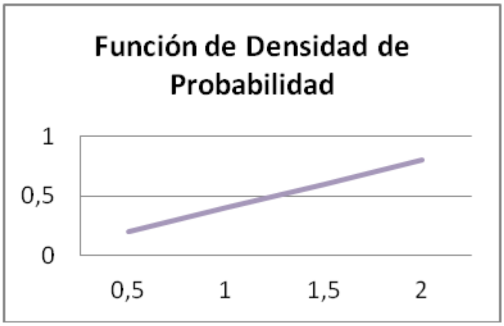


Ilustración 6.8: Densidad de probabilidad

b) Empleamos la expresión de función de distribución acumulada:

$$F(x) = P(X \leq x) = \int_{\infty}^{-\infty} f(y)\, dy$$

En nuestro caso, el intervalo que tenemos que realizar cálculos es $0 \leq Y \leq 2$:

$$F(y) = \int_0^y (\frac{1}{8} + \frac{3t}{8})\, dt = \frac{1}{8} t \Big|_0^y + \frac{3}{16} t^2 \Big|_0^y = \frac{1}{8} y + \frac{3}{16} y^2 = \frac{y}{16}(2 + 3y)$$

Si ***y < 0***, la función de distribución es igual a ***0***, y si ***y > 2***, la función de distribución es **1**, por consiguiente, la función de distribución queda como se muestra a continuación:

$$f(y) = \begin{cases} 0; & y < 0 \\ \frac{y}{16}(2 + 3y) & 0 \leq y \leq 2 \\ 1 & y > 2 \end{cases}$$

c) Para calcular la probabilidad de este apartado, empleamos la función de distribución acumulada tal y como se muestra a continuación:

P ($1 \leq Y \leq 1.5$) = F (1.5) – F (1) = (1.5/16) (2 + 3·1.5) - (1/16) (2 + 3) = 19/64

d) Para calcular la probabilidad que nos piden en este apartado, al igual que el anterior, hacemos uso de la función de distribución acumulada:

$$P(Y > 1) = 1 - P(Y \leq 1) = 1 - F(1) = 1 - (1/16)(2 + 3) = 11/16$$

6.3.4 Distribución normal

La distribución de probabilidad normal contínua es muy importante en el campo de la estadística, su gráfica, denominada curva normal, es la curva con la cual se representa de forma aproximada muchos fenómenos que ocurren en la industria, la naturaleza y la investigación. La distribución normal muchas veces se denomina **distribución Gaussiana**.

La ecuación para la distribución de probabilidad de la variable normal obedece a los dos parámetros **μ** y **σ**, su media y su desviación estándar respectivamente, por ello expresamos los valores de la densidad de **X** por ***n(x; μ, σ).*** La densidad de la variable aleatoria normal ***X***, con media ***μ*** y varianza ***σ²***

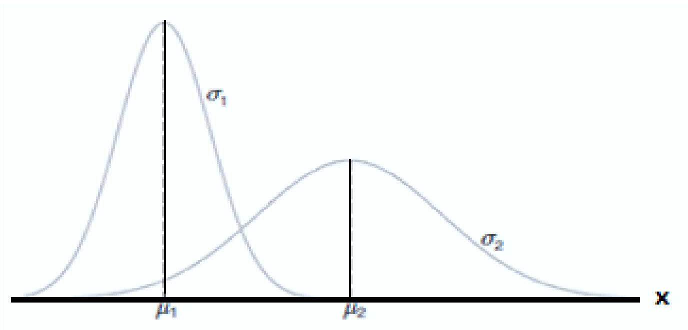


Ilustración 6.4 Representación de la Curva Normal
Curvas Normales con $\mu_1 < \mu_2$; $\sigma_1 < \sigma_2$

[6.28] $$\boldsymbol{n(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}}\left(e^{-\frac{1}{2\sigma^2} * (x-\mu)^2}\right)};\ -\infty \leq \boldsymbol{x} \leq \infty$$

La media y la varianza de ***n (x; μ, σ)*** son μ y ***σ²***, respectivamente. Por lo tanto, la desviación estándar es ***σ***.

$$E(x - \mu) = \int_{-\infty}^{\infty} \frac{x - \mu}{\sqrt{2\pi\sigma}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Al establecer que z = (x – μ)/σ y dx = σ dz, obtenemos

$$E(x-\mu) = \frac{1}{\sqrt{2\pi\sigma}}\int_{-\infty}^{\infty} ze^{-\frac{1}{2}z^2}\,dz = 0$$

Condicionalmente la integral anterior es una función impar de z concluimos que

E(X) = μ

La varianza de la distribución normal es dada por:

$$E[(x-\mu)^2] = \frac{1}{\sqrt{2\pi\sigma}}\int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,dx$$

De nuevo, al establecer que z = (x – μ)/σ y dx = σ dz, obtenemos

$$E[(x-\mu)^2] = \frac{\sigma^2}{\sqrt{2\pi\sigma}}\int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}}\,dz$$

Al integrar por partes con ***u = z*** y ***dv = ze−z ² /2 dz*** de modo que **du = dz** y ***v =−e***$^{-z/2}$, encontramos que:

$$E[(x-\mu)^2] = \frac{\sigma^2}{\sqrt{2\pi\sigma}}\left(-z^2 e^{-\frac{z^2}{2}} + \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}\,dz\right) = \sigma^2(0+1) = \sigma^2$$

Se puede convertir todas las informaciones de cualquier variable aleatoria normal X en un conjunto de análisis de una variable aleatoria normal Z con media 0 y varianza 1. Esto se puede realizase mediante la transformación

[6.29] $$Z = \frac{X-\mu}{\sigma}$$

La distribución de una variable aleatoria normal con media 0 y varianza 1 se llama **distribución normal estándar.**

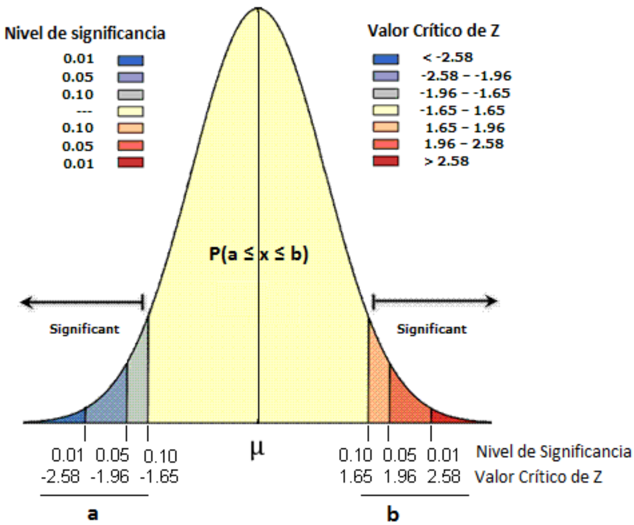


Ilustración 6.5 Sesgo de Probabilidad Gaussiana

Gráficamente la distribución normal estándar da la probabilidad de la variable X que dependerá del área del perímetro sombreado en la figura y para calcularla se utilizara una tabla, que se encuentra en el Apéndice.

Características de la Distribución Uniforme Contínua

- Posee simetría, respecto de **x = 0**, ya que:

$$f(-x) = f(x)$$

- Alcanza su punto máximo en **x = 0**, donde vale:
  - $f(0) = \frac{1}{\sqrt{2\pi}}$
- La simetría permite que sea creciente para **x < 0** y viceversa (decreciente) para **x > 0**.
- Los puntos inflexión en las abscisas para la función son de 1 y -1.

Ejemplo 6.25: Según el departamento de meteorología, la temperatura durante el mes de septiembre está distribuida de manera normal, con media 18,7ºC y desviación standard 5ºC. Hallar la probabilidad de que la temperatura durante septiembre esté por debajo de 21ºC.

Resolución

$$\mu = 18{,}7°C \quad \sigma = 5°C \quad X = 21°C$$

$$Z = \frac{X - \mu}{\sigma} = \frac{21 - 18{,}7}{5} = \frac{2{,}3}{5} = 0{,}46$$

Con el valor de Z = 0,46, se localiza en la tabla y tenemos que la probabilidad es de 0,6772 (67,72%)

0,46

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	[illegible]	0,8264	[illegible]	0,8315	0,8340	[illegible]	0,8389
	0.00	0.01	[illegible]	[illegible]	0.04		0.06	0.07		0.09

Ilustración 6.6 Representación de Valores en la Tabla

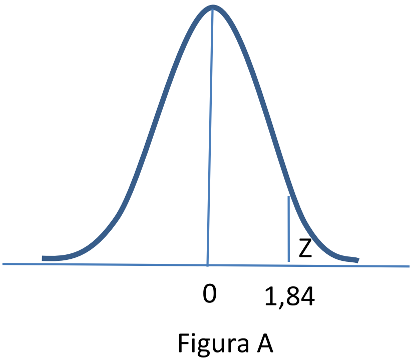
Ejemplo 6.26: Grafique según una distribución normal estándar, calcule el área bajo la curva que se limita
a) a la derecha de z = 1.84, y
b) entre z = –1.97 y z = 0.86.

a) El área en la figura **A**, a la derecha de z = 1.84 es***:***

- ***Z = 1 – 0.9671 = 0.0329.***

b) El área en la figura **B** entre z = –1.97 y z = 0.86 es***:***

- ***Z = 0.8051 – 0.0244 = 0.7807****.*



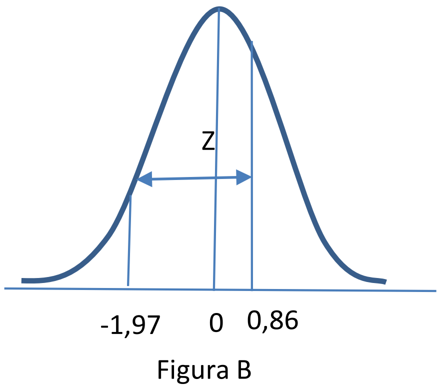


Ilustración 6.7 Área Bajo la Curva

Ejemplo 6.27: En una ciudad de Quito se calcula que la temperatura máxima en el mes de junio sigue una distribución normal, con media 23° y desviación estándar típica 5°. Hallar el número de días del mes en los que se prevé alcanzar máximas entre 21° y 27°.

$$\text{P }(21 < x < 27) = p(\frac{21-23}{5} < Z \leq \frac{27-23}{5})$$
$$p(-0{,}4 < Z \leq 0.8) = p(Z \leq 0{,}8) - [1 - p(z \leq 0{,}4)]$$

Valores de la Tabla Z = 0.7881 – (1-0.6554) = 0.4425* 30 días = 13

6.3.5 Distribución T Student

En unidades anteriores se empleó el uso de la distribución z, la cual se manipula los tamaños de las muestras que fueran mayores o ≥30, o en muestras más pequeñas si la distribución o las distribuciones del cual proviene la muestra o las muestras son normales.

La teoría de pequeñas muestras también se le llama ***teoría exacta del muestreo***, ya que también la podemos utilizar con muestras aleatorias de tamaño grande.

También se verá un nuevo concepto necesario para poder utilizar la distribución mencionada. Este concepto es "grados de libertad" que sirve para delimitar grados de libertad se forjará en referencia a la varianza muestral.

Características de las distribuciones T Student
Las características básicas de la distribución Student son:

- Cada curva t tiene forma de campana con centro en 0.
- Cada curva t, está más dispersa que la curva normal estándar z.
- A medida que n aumenta, la dispersión de la curva t proporcionada disminuye.
- A medida que n $\rightarrow \infty$, la sucesión de curvas t se acerca a la curva normal estándar, por lo que la curva z toma a veces el seudónimo de curva t con Gl = ∞

La distribución de la variable aleatoria t está dada por:

[6.30] $$\beta(t) = \frac{\Gamma\left[\frac{(v+1)}{2}\right]}{\Gamma\left(\frac{v}{2}\right)\sqrt{u}}\left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}} \; ; \; -\infty < t < \infty$$

Con grados de libertad

[6.31] $$t = \frac{X - \mu}{\frac{S}{\sqrt{n}}}$$

La distribución contínua t es difiere de la distribución Z en que la varianza de t depende del tamaño que tiene la muestra y constantemente es mayor a uno. Solo cuando el tamaño de la muestra se expande a infinito las dos distribuciones serán las equivalentes, se acostumbra representar con $\mathbf{t_\propto}$ el valor de **t** por arriba del cual se encuentra un área igual a **∝**. Como la distribución **t** es simétrica alrededor de una media de cero, tenemos $\mathbf{t_{1-\propto}}$; es decir, el valor **t** que deja un área de **1-∝** a la derecha y por tanto un área de **∝** a la izquierda, es igual al valor **t** negativo que deja un área de **∝** en la cola derecha de la distribución.

Intervalo de Confianza para μ; con σ desconocida
Si X y S son la media y la desviación estándar de una muestra aleatoria de una población normal con varianza σ2, desconocida, un intervalo de confianza de (1 - **∝**) 100% para μ es:

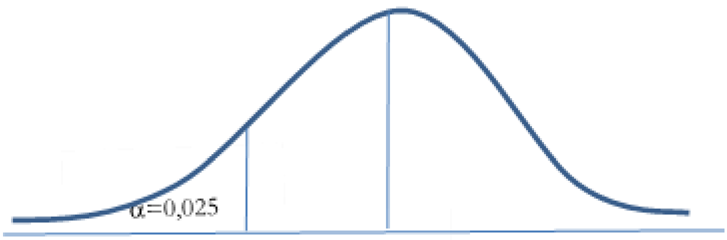
[6.32] $$X - t_{\propto/2}\frac{S}{\sqrt{n}} < \mu < X + t_{\propto/2}\frac{S}{\sqrt{n}}$$

Donde $t_{\propto/2}$ es el valor t con $v = n\text{-}1$ grados de libertad, que deja un área de $\propto/2$ a la derecha.

Se hace una distinción entre los casos de σ conocida y σ² desconocida al calcular las estimaciones del intervalo de confianza. Un ejemplo de este caso se presenta en el ejemplo número 3.

Ejemplo 6.28: El valor t con v= 14 GL que resigna un área de 0.025 a la izquierda, y por tanto un área de 0.975 a la derecha, es:(Hallar valores en la tabla t)

$$t_{0.975} = -t_{0.025} = -2.145$$



$t_{0.975} = -t_{0.025} = -2.145$

Ilustración 6.9: Valores de la Distribución T

Ejemplo 6.29: Un Agente de Certificación de Calidad afirma que el rendimiento medio de la productora de alimentos donde cierto proceso en lotes es 500 gramos por milímetro de materia procesada. Para verificar esta aseveración escoge una muestra de 25 lotes cada mes. Donde el valor de t calculado se halla entre –t0.05 y t0.05, queda percibida con su afirmación. ¿Qué solución extraería de una muestra que posee una media de 518 gramos por milímetro y una desviación estándar de 40 gramos? Suponga que la distribución es aproximadamente normal.
De la tabla encontramos que t0.05 para 24 grados de libertad es de 1.711. Por tanto, el agente queda satisfecho con esta afirmación si una muestra de 25 lotes rinde un valor t entre –1.711 y 1.711.

Calculando el valor de t:

$$t = \frac{X - \mu}{\frac{S}{\sqrt{n}}} = \frac{518 - 500}{\frac{40}{\sqrt{25}}} = 2,25$$

Es un valor por arriba de 1.711. Si se anhela lograr la probabilidad de conseguir un valor de t con 24 grados de libertad **≥2.25** se busca en la tabla y es aproximadamente de 0.02.

Ejemplo 6.30: El contenido de siete envases con químicos que contienen de ácido nítrico son 9.8; 10.2; 10.4; 9.8; 10.0; 10.2 y 9.6 litros. Encuentre un intervalo de confianza del 95% para la media de todos los envases si se supone una distribución aproximadamente normal. Donde X = 10 y S = 0.283.

En la tabla se encuentra que t0.025=2.447 con 6 grados de libertad, de aquí, el intervalo de confianza de 95% para μ es:

$$10 - (2{,}477)(\frac{0{,}283}{\sqrt{7}} < \ \mu < 10 + (2{,}477)\ (\frac{0{,}283}{\sqrt{7}})$$
$$9{,}47 < \ \mu < 10{,}26$$

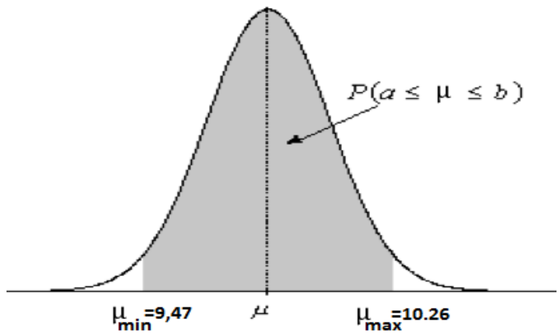


Ilustración 6.10: Representación de la Media

Con un nivel de confianza del 95% , se está al tanto que el promedio del contenido de los embaces está entre 9.47 y 10.26 litros.

6.3.6 Distribución Chi cuadrado (X^2) o de Pearson

Para el caso especial de la distribución Chi cuadrado se obtiene al permitir que $\alpha = v/2$ y $\beta = 2$, donde n es un entero positivo. La distribución posee un solo parámetro, **n**, denominado grados de libertad. La Distribución Chi Cuadrada es aquella variable aleatoria continúa **X**, con **n** grados de libertad y su función de densidad es dada por:

[6.33] $$f(x;n) = \begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{n/2-1} e^{-x/2}, & 0 < x < \infty \\ 0, & en\ otros\ casos \end{cases}$$

Donde n es un entero positivo.

La distribución contínua Chi cuadrado desempeña un título fundamental en la inferencia estadística, contiene una concentración considerable de mecanismos evaluadores tanto en la metodología como en la teoría. La distribución Chi Cuadrado es un mecanismo significativo de la prueba estadística de hipótesis y de la estimación estadística.

La distribución Chi cuadrada es demostrada a continuación:

[6.34] **Media:** $\mu = n$
[6.35] **Varianza:** $\sigma 2 = 2n$

Características de las Distribuciones Chi cuadrado (X^2) o de Pearson
Las características básicas de la distribución de Pearson (Chi cuadrado) son:

- No posee o adopta valores menores de cero.
- A medida que aumentan los grados de libertad se va aproximando a la distribución normal.
- Es asimétrica positiva.
- Para muestras >100 podemos aproximar mediante una distibución normal.

Ejemplo 6.31: La compañía tecnológica SONY ha desarrollado una nueva batería para celulares. En promedio, la batería de litio dura 60 minutos por carga. La desviación estándar es de 4 minutos. Supongamos que el departamento de producción corre una prueba de control de calidad. Ellos seleccionan 7 baterías al azar. La desviación estándar de las baterías seleccionadas es de 6 minutos. ¿Qué valor de la estadística Chi Cuadrada tenemos para esta prueba?

Para calcular la estadística Chi Cuadrada, usamos los valores en la ecuación para χ^2

$$X^2 = \frac{(n-\ 1)S^2}{\sigma^2} = \frac{(7-1)6^2}{4^2} = 13{,}5$$

Dado el valor, podemos determinar la probabilidad acumulada de Chi Cuadrada. Para esto, usamos una tabla de la estadística Chi Cuadrada **($\chi 2$)** con los valores de GL **(6)** y de Chi- cuadrada **(13.5)** o usamos alguna herramienta para el cálculo de Chi Cuadrada. De cualquiera de los dos obtenemos el valor de: **0,96**. Lo anterior significa que la probabilidad de que la desviación estándar sea mayor a **6** minutos es de **1 – 0,96** o sea **0,04** (muy pequeña).

6.4 Ejercicios propuestos

1.- Supongamos que la probabilidad de poseer una unidad imperfecta en una línea ensambladora es de 0.05. Si el acumulado de unidades terminadas forma parte de un conjunto de ensayos independientes:

a) ¿cuál es la probabilidad de que entre diez unidades dos se hallen imperfectas?
b) ¿y de que a lo máximo dos se encuentren imperfectas?
c) ¿cuál es la probabilidad de que por lo menos una se encuentre imperfecta?

2.- El administrador de un restaurante que proporciona servicio mediante reservas sabe, por costumbre, que el 20% de los clientes que reservan una mesa no estarán. Si el establecimiento acepta 25 reservas pero únicamente dispone de 20 mesas, ¿cuál es la probabilidad de que a todas las personas que pidieron una reserva, asistan al restaurante y se les dé una mesa?

3.- En un laboratorio se está realizando un experimento medio ambiental donde una muestra de aire tiene 10% de probabilidad de contener una molécula rara particular. Suponga que las muestras son independientes con relación a la representación de la molécula rara. Halle la probabilidad de que en las subsecuentes 18 muestras, puntualmente 2 contengan la molécula rara.

4.- Se lanza un dado equilibrado originando el espacio equis probable S= {1, 2, 3, 4, 5, 6}. Sea X el doble del número que aparece. Encuentre la distribución f, la media μ x, la varianza σx2 y la desviación estándar σx de X.

5.- Un examen consta de 200 preguntas de verdadero o falso, para un sujeto que respondiese al azar. Cuál sería la probabilidad de que acertase:

- 50 preguntas o menos.
- Más de 50 y menos de 100.
- Más de 120 preguntas.

6.- ¿Cuántos números de 4 dígitos se pueden formar con las cifras 0, 1, 2, 3, 4, 5, 6, 7, 8 ,9?

- Permitiendo repeticiones.
- Sin repeticiones.
- Si el último dígito ha de ser 0 y no se permiten repeticiones.

7.- Utilizando las fórmula de la distribución binomial B (6; 0,15) calcular las probabilidades:
a) P (X = 2) b) P (X > 0) c) P (X ≤ 3)
d) Calcule la esperanza matemática, varianza, moda y asimetría de la distribución B (6; 0,15).

8.- En el conjunto de Pediatría de un Hospital fueron atendidos 200 niños en el último mes. Conociendo que 130 de ellos mostraron afecciones gripales, ya que se eligieron al azar 9 historiales clínicos, calcule la probabilidad de que:
a) todos padeciesen gripe
b) alguno hubiera padecido gripe.

9.- Una empresa donde se ensamblan IPhone, hay en una mesa con 20 componentes (circuitos) de los cuales 6 están deteriorados. Al comienzo llega el ensamblador y recoge 8 circuitos y más tarde llega otro ensamblador y se lleva los restantes. Hallar la probabilidad que solamente uno de ensambladores se haya llevado todos los chips defectuosos.

10.- Para evitar el sistema de detección de tráfico en la aduana, un contrabandista ha colocado 6 tabletas de narcótico en un frasco que contiene 9 píldoras de medicina que son de igual apariencia. Si los oficiales de supervisión de la aduana seleccionan 3 tabletas al azar para analizar las tabletas.

a) ¿Cuál es la posibilidad de que el contrabandista sea detenido por posesión de narcóticos?
b) ¿Cuál es la probabilidad de que no sea detenido?

11.- Un cohete tiene 3 computadoras idénticas. Solamente una de ellas se utiliza para controlar la dirección del cohete, las otras 2 computadoras son de reserva, redundantes, por si acaso la primera sufra algún desperfecto al momento del lanzamiento. Durante una hora de maniobras la probabilidad de que falle la computadora es 0.0005.

a.- ¿Cuál es el tiempo promedio de fallo de las tres computadoras?
b.- ¿Cuál es la probabilidad de que las 3 fallen durante un vuelo de 5 horas?

12.- De acuerdo con la ciencia de la genética, un genetista ha realizado un experimento de cierto cruce en ganado vacuno donde resultará en una descendencia roja, negra y blanca en la relación 8: 4: 4. Encuentre la probabilidad de que entre 8 descendientes:
a) 5 sean rojos, 2 negros y un blanco,
b) 3 sean rojos y 2 sean negros.

13.- Un grupo de 12 personas decide reunirse en cierta ciudad. La probabilidad de que una persona llegue a la ciudad en un avión, coche, tren o autobús es, respectivamente .3, .4, .1 y .2 ¿Cuál es la probabilidad de que de las 12 personas, 3 lleguen en avión, 5 en coche, 2 en tren y 2 en autobús?

14.- El director técnico de un equipo de fútbol opina que los jugadores A, B y C tienen las mismas aptitudes para ser titulares del equipo en la posición de delanteros. Entonces, determina que los tres jueguen el mismo número de minutos en cada partido. Se conoce que el 40% de los goles son de C, mientras que A y B consiguen un 30% de goles. ¿Cuál sería la probabilidad de que en un partido con 9 goles, A consiguiera marcar dos, B marcar tres y C marcar cuatro?

15.- En la hora pico de la mañana un semáforo (que se encuentra en una calle transitada) pasa una media de 8 autos por minuto. ¿Cuál es la probabilidad de que pasen 16 autos en una interrupción de 2 minutos?

16.- En la zona protegida de bosque cercana a un punto kilométrico de una autopista hay una población de armadillos. Se ha elaborado un estudio por el que se ha concluido que los armadillos de la reserva cruzan la autopista a razón de 1 armadillo cada 20 minutos. ¿La posibilidad de que 4 armadillos crucen la autopista en una hora es?

17.- El promedio de vida de los pobladores del valle de Vilcabamba es de 100 años, con una varianza de 0,35. Se realiza un experimento en la pequeña ciudad de 10.000 habitantes:
a) ¿Cuántas personas superarán previsiblemente los 105 años?
b) ¿Cuántos vivirán menos de 90 años?

18.- Un biólogo desea medir la concentración que cierto contaminante ambiental que se encuentra distribuida de modo uniforme en el intervalo de 0 a 20 pares de millón. Si se estima un nivel de toxicidad con una concentración de 8 o más. ¿Cuál es la expectativa de que al tomar una muestra, tenga una densidad tóxica? ¿Probabilidad de que la concentración toxica sea exactamente 10?

19.- Una agrupación de variable aleatoria se distribuye uniformemente en el intervalo [0, 100].

- Si su valor se encuentra comprendido entre 20 y 35, calcule su probabilidad.
- Determine su esperanza matemática y varianza.

20.- Determine la función de distribución, esperanza matemática y varianza de la variable aleatoria contínua X que se distribuye uniformemente en el intervalo [2 , 4] y para dicha variable, calcule las probabilidades: $Pr(X \geq 3)$ $Pr(1{,}25 < X \leq 2{,}05)$.

21.- Si X es una variable normal aleatoria de una distribución N (μ, σ), encontrar:

- $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma)$

22.- En la ciudad de Guayaquil una de cada tres hogares posee el servicio de internet. Si se eligen al azar 90 hogares, calcular la probabilidad de que entre estos haya por lo menos 30 tengan internet.

23.- En un experimento de laboratorio se conoce que el tiempo de incubación de cierta especie en a una determinada cámara es una variable aleatoria que sigue una distribución Gamma (5, 10), ¿cuál es la probabilidad de que la especie encubada no supere las 60 semanas de vida?

24.- Un componente que pertenece al sistema eléctrico de una máquina automatiza de una empresa, falla una vez cada 5 horas, entonces: ¿cuál sería el tiempo medio que pasa hasta que fallan dos componentes? ¿Qué posibilidad hay de que sucedan 12 horas antes que fallen los dos componentes?

25.- En una cafetería el tiempo que transcurre para una persona para ser atendida es una tiene una distribución exponencial con una media de 4 minutos. ¿Cuál sería la posibilidad de que una persona se le brinde atención antes de que pasen 3 minutos en al menos 4 de los 6 días siguientes?

26.- Suponga que una empresa ensambla computadoras que poseen cierto tipo de circuito cuyo tiempo de falla en años está dado por una distribución exponencial, con tiempo promedio de falla $\boldsymbol{\lambda} = 5$. Sí 5 de estos circuitos se instalan en diferentes computadoras, ¿cuál es la expectativa de que al menos 2 computadores continúen funcionando después de 8 años?

27.- Las calificaciones de los estudiantes de un curso de la carrera de diseño se componen de una distribución de variables aleatorias, con promedio de 50, pese a que no se conoce la desviación. Se tomó una media de 9 estudiantes hallando una varianza de 36 y un promedio de 52. ¿Cuál es la probabilidad de que el promedio:

- Sea mayor de 54?
- Sea menor que 54?
- Esté comprendido entre 48 y 52 puntos?

28.- Un fábrica de baterías asegura que sus baterías de litio perdurará un promedio de 500 horas de trabajo. Para mantener este promedio el productor verifica 25 baterías cada mes. Si el valor calculado queda entre $-t$ 0.05 y t 0.05, el productor se encuentra tranquilo con esta afirmación. ¿Qué desenlace deberá él sacar de una muestra de 25 baterías cuya duración fue (horas)?

520	521	511	513	510
513	522	500	521	495
496	488	500	502	512
510	510	475	505	521
506	503	487	493	500

29.- En una máquina expendedora de snacks existen 4 canales que expiden el mismo tipo de bebida. La consecuente tabla exhibe el número de snacks vendidos en cada uno de los 4 canales durante una semana. Contrastar la hipótesis de que los canales son seleccionados al azar a un nivel de significación del 5%.

Canal	**Número de snacks consumidas**
1	13
2	22
3	18
4	17

30.- Una pequeña población que posee una distribución aproximadamente normal adquiere una varianza definida de 0.8. Hallar la probabilidad que una muestra aleatoria de n = 6 asuma una varianza mayor o igual a 1,2.

6.5 Ecuaciones introducidas en el capítulo 6

Distribución uniforme

La distribución uniforme tiene exactamente la misma probabilidad para todos sus elementos, es en la cual la variable aleatoria asume cada uno de sus valores con idéntica probabilidad.

[6.1] $$F(x) = \frac{1}{x}\text{; Donde } x = x_1, x_2, x_3 \ldots\ldots\ldots x_n$$

[6.2] $$\text{Media: } \mu = E(x) = \frac{\sum_{i=1}^{k} xi\, f(x)}{k}$$

[6.3] $$\text{Varianza: } \sigma^2 = E(x)^2 - [E(x)]^2$$

Distribución de Bernoulli

La distribución de Bernoulli es un caso especial de la distribución binomial, donde *n* es el número de experimentos, *x* número de existo, θ probabilidad de éxito y $(1-\theta)$ probabilidad de fracaso.

[6.4] $$b\ (x) = p * q^{\ 1-p}$$

Distribución binomial

La distribución de probabilidad binomial y de Bernoulli es una variable aleatoria discreta que se presenta con mucha frecuencia.

[6.5] $$B\ (x; n, p) = \binom{n}{x} p^x q^{n-x},\ x = 0, 1, 2 \ldots n.$$

[6.6] **Media:** $\mu = n*p$

[6.7] **Varianza:** $\sigma^2 = n*p*q$

Distribución binomial negativa

Si pruebas independientes repetidas pueden dar como consecuencia un éxito con probabilidad *p* y un fracaso con probabilidad $q = 1 - p$, en aquella distribución de probabilidad de la variable aleatoria *X*, el número de pruebas en el que ocurre el k-ésimo éxito, es:

[6.8] $$B^*\ (x; k, p) = \binom{x-1}{k-1}\ p^k\ q^{\ x\text{-}k};\ k = k,\ k+1,\ k+2\ldots.$$

[6.9] **Esperanza:** $\mu = \frac{(k-q)}{p}$

[6.10] **Varianza:** $\mu = \frac{(k-q)}{p^2}$

Distribución multinomial

La distribución del experimento binomial se transforma en el caso de en un experimento multinomial si cada prueba posee más de dos resultados posibles.

[6.11] $$f(x_1, x_2 \ldots . x_k; p_1, p_2 \ldots p_k) = \binom{n}{x_1, x_2 \ldots . x_k} {p_1}^{x1} * {p_2}^{x2} \ldots . {p_k}^{xk}$$

$$\sum_{x=1}^{k} x_i = n \ y \sum_{x=1}^{k} p_i = 1$$

Distribución geométrica

Pruebas independientes frecuentadas pueden obtener como consecuencia un éxito con probabilidad p y un fracaso con probabilidad $q = 1 - p$, entonces la distribución de probabilidad de la variable aleatoria X, el número de la prueba en el que ocurre el primer éxito, es:

[6.12] $$G(x; p) = p * q^{x-1},\ x = 1, 2, 3, \ldots n$$

[6.13] $$\textbf{Media: } \mu = \frac{1}{p}$$

[6.14] $$\textbf{Varianza: } \sigma^2 = \frac{1-p}{p^2}$$

Distribución hipergeométrica

La distribución de probabilidad de la variable aleatoria Hipergeométrica X, el número de éxitos en una muestra aleatoria de tamaño n que se selecciona de N artículos, en los que k se denomina éxito y N – k fracaso.

[6.15] $$h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}\ ;\max\left(0, n - (N-k)\right) \leq x \leq \min(n, k)$$

[6.16] $$\textbf{Media: } \mu = \frac{nk}{N}$$

[6.17] $$\textbf{Varianza: } \sigma^2 = \frac{N-n}{N-1} * n * \frac{k}{N} * \left(1 - \frac{k}{N}\right)$$

Distribución de Poisson

La distribución de Poisson es una distribución de probabilidad discreta que se emplea a los procesos de algún evento durante un intervalo específico de medida cuantitativa.

[6.18] $$f(x = k) = \left(\frac{e^{-\mu}(\mu)^x}{x!}\right)$$

Si dentro de las ocurrencias de un suceso μ se presentan intervalos, μ pasa a determinarse por **λ**t, donde **λ** representa la razón media por unidad y **t** el número de veces que ocurre, por lo cual la función de Poisson se puede también representar:

[6.18 A] $$P(x;\ \lambda t) = \left(\frac{e^{-\lambda t}(\lambda t)^x}{x!}\right)$$

Distribución uniforme contínua

La parte de distribución contínua más sencilla de la estadística es la distribución uniforme contínua.

[6.19] $$f(x; A, B) = \begin{cases} \frac{1}{B-A}, & A \leq x \leq B \\ 0, & en\ otros\ casos \end{cases}$$

[6.20] **Media:** $\mu = \frac{A+B}{2}$

[6.21] **Varianza:** $\sigma^2 = \frac{(B-A)^2}{12}$

Distribución Gamma

La distribución Gamma se trata de una parte de distribuciones que provee un modelo adecuado para histogramas que presentan cierto tipo de asimetría. Antes de presentar a las variables aleatorias con distribución Gamma, es necesario recordar cómo se define la función Gamma o factorial, la cual cumple un rol importante en muchas ramas de la Matemática.

[6.22] $$f(x; \propto, \beta) = \begin{cases} \frac{1}{\beta^{\propto}\,\Gamma(\propto)}\, x^{\propto -1} e^{-x/\beta}, & 0 > x \\ 0, & en\ otros\ casos \end{cases}$$

[6.23] **Media:** $\frac{p}{a}$

[6.24] **Varianza:** $\frac{p}{a^2}$

Distribución exponencial

La distribución exponencial es un caso especial de la distribución gamma. La relación entre la gamma y la exponencial permite que la distribución gamma se utilice en tipos similares de problemas.

[6.25] $$f(x; \beta) = \begin{cases} \frac{1}{\beta}\, e^{-x/\beta}, & 0 > x \\ 0, & en\ otros\ casos \end{cases}$$ **Donde $\beta > 0$**

[6.26] $$f(x) = \int_0^x \propto e^{-\propto x} dx = [-e^{-\propto x}] = 1 - e^{-\propto x}$$

[6.27] **Media:** $\mu = \beta$

Varianza: $\sigma^2 = \beta^2$

Distribución normal

La distribución de probabilidad normal contínua es muy importante en el campo de la estadística, su gráfica, denominada curva normal, es la curva con la cual se representa de forma aproximada muchos fenómenos que ocurren en la industria, la naturaleza y la investigación. La distribución normal muchas veces se denomina distribución Gaussiana.

[6.28] $$\boldsymbol{n(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}}\left(e^{-\frac{1}{2\sigma^2} * (x-\mu)^2}\right);\ -\infty \leq x \leq \infty}$$

Distribución normal estándar

Se puede convertir todas las informaciones de cualquier variable aleatoria normal X en un conjunto de análisis de una variable aleatoria normal Z con media 0 y varianza 1. Esto se puede realizase mediante la transformación

[6.29] $$\boldsymbol{Z = \frac{X - \mu}{\sigma}}$$

Distribución T student

En unidades anteriores se empleó el uso de la distribución z, la cual se manipula los tamaños de las muestras que fueran mayores o ≥30, o en muestras más pequeñas si la distribución o las distribuciones del cual proviene la muestra o las muestras son normales.

[6.30] $$\boldsymbol{\beta(t) = \frac{\Gamma\left[\frac{(v+1)}{2}\right]}{\Gamma\left(\frac{v}{2}\right)\sqrt{u}}\left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}};\ -\infty < t < \infty}$$

Con grados de libertad

[6.31] $$t = \frac{\boldsymbol{X - \mu}}{\frac{\boldsymbol{S}}{\sqrt{\boldsymbol{n}}}}$$

Intervalo de Confianza para μ; con σ desconocida

Si X y S son la media y la desviación estándar de una muestra aleatoria de una población normal con varianza σ2, desconocida, un intervalo de confianza de (1 - ∝) 100% para μ es:

[6.32] $$\boldsymbol{X - t_{\propto/2}\frac{S}{\sqrt{n}} < \mu < X + t_{\propto/2}\frac{S}{\sqrt{n}}}$$

Distribución Chi cuadrado (X2)

Para el caso especial de la distribución Chi Cuadrado se obtiene al permitir que α = v/2 y β = 2, donde *n* es un entero positivo. La distribución posee un solo parámetro, n, denominado grados de libertad. La Distribución Chi Cuadrada es aquella variable aleatoria continúa X, con n grados de libertad y su función de densidad es dada por:

[6.33] $$f(\boldsymbol{x};\boldsymbol{n}) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}\, x^{n/2-1} e^{-x/2}, & 0 < x < \infty \\ 0, & \textit{en otros casos} \end{cases}$$

[6.34] **Media:** $\mu = n$

[6.35] **Varianza:** $\sigma^2 = 2n$

7. Apéndice

7.1 Tabla I

Tabla de valores de probabilidad acumulada (para la Distribución Normal Estándar (Walpole, 1999)

	0	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
-3	0,0013	0,0010	0,0007	0,0005	0,0003	0,0002	0,0002	0,0001	0,0001	0,0000
-2,9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879

	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3	0,9987	0,9990	0,9993	0,9995	0,9997	0,9998	0,9998	0,9999	0,9999	1,0000

7.2 Tabla II

Distribución Normal: En los ejes están los valores de z y la tabla muestra el área del eje central a la derecha.

z	**0,00**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

La tabla posee los valores del eje central a la derecha. En la función de **Excel =Distr.norm.estand(z)** corresponde el resultado es el área de (-) infinito al valor de z, por lo que al resultado es necesario restarle 0.5 (Levin, 2010)

7.3 Tabla III

Distribución t de student: En los renglones se ubican los grados de libertad. En las columnas son las áreas más usadas. La tabla son los valores del eje x (t - área)

G.l.	$t_{0.1}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$
1	3,0777	6,3138	12,7062	31,8205	63,6567
2	1,8856	2,9200	4,3027	6,9646	9,9248
3	1,6377	2,3534	3,1824	4,5407	5,8409
4	1,5332	2,1318	2,7764	3,7469	4,6041
5	1,4759	2,0150	2,5706	3,3649	4,0321
6	1,4398	1,9432	2,4469	3,1427	3,7074
7	1,4149	1,8946	2,3646	2,9980	3,4995
8	1,3968	1,8595	2,3060	2,8965	3,3554
9	1,3830	1,8331	2,2622	2,8214	3,2498
10	1,3722	1,8125	2,2281	2,7638	3,1693
11	1,3634	1,7959	2,2010	2,7181	3,1058
12	1,3562	1,7823	2,1788	2,6810	3,0545
13	1,3502	1,7709	2,1604	2,6503	3,0123
14	1,3450	1,7613	2,1448	2,6245	2,9768
15	1,3406	1,7531	2,1314	2,6025	2,9467
16	1,3368	1,7459	2,1199	2,5835	2,9208
17	1,3334	1,7396	2,1098	2,5669	2,8982
18	1,3304	1,7341	2,1009	2,5524	2,8784
19	1,3277	1,7291	2,0930	2,5395	2,8609
20	1,3253	1,7247	2,0860	2,5280	2,8453
21	1,3232	1,7207	2,0796	2,5176	2,8314
22	1,3212	1,7171	2,0739	2,5083	2,8188
23	1,3195	1,7139	2,0687	2,4999	2,8073
24	1,3178	1,7109	2,0639	2,4922	2,7969
25	1,3163	1,7081	2,0595	2,4851	2,7874
26	1,3150	1,7056	2,0555	2,4786	2,7787
27	1,3137	1,7033	2,0518	2,4727	2,7707
28	1,3125	1,7011	2,0484	2,4671	2,7633
29	1,3114	1,6991	2,0452	2,4620	2,7564

En la tabla se elige la columna con el área deseada, después se localiza el valor del eje **x**. A esta táctica se llama función inversa (en la forma directa se tiene el valor de **x** y la tabla proporciona el área). En Excel se usa la función**:=Distr.t.inv(prob,gl)** (Webster Allen, 2008)

Pero brinda el área "a dos colas", es decir, que se busca un área de **0.05**, significa que **0.025** esta de un lado y del otro. Para trabajar la cota con una cola, como la tabla, necesario multiplicar por 2 la probabilidad. t 0.05,6 se busca como **=distr.t.inv(0.1,6)**

7.4 Tabla IV

Distribución **(X²)** Chi-Cuadrado.

0,001	**0,005**	**0,01**	**0,02**	**0,025**	**0,03**	**0,04**	**0,05**	**0,10**	**0,15**	**0,20**	**0,25**	**0,30**	**0,35**	**0,40**	**G.L.**
10,828	7,879	6,635	5,412	5,024	4,709	4,218	3,841	2,706	2,072	1,642	1,323	1,074	0,873	0,708	**1**
13,816	10,597	9,210	7,824	7,378	7,013	6,438	5,991	4,605	3,794	3,219	2,773	2,408	2,100	1,833	**2**
16,266	12,838	11,345	9,837	9,348	8,947	8,311	7,815	6,251	5,317	4,642	4,108	3,665	3,283	2,946	**3**
18,467	14,860	13,277	11,668	11,143	10,712	10,026	9,488	7,779	6,745	5,989	5,385	4,878	4,438	4,045	**4**
20,515	16,750	15,086	13,388	12,833	12,375	11,644	11,070	9,236	8,115	7,289	6,626	6,064	5,573	5,132	**5**
22,458	18,548	16,812	15,033	14,449	13,968	13,198	12,592	10,645	9,446	8,558	7,841	7,231	6,695	6,211	**6**
24,322	20,278	18,475	16,622	16,013	15,509	14,703	14,067	12,017	10,748	9,803	9,037	8,383	7,806	7,283	**7**
26,124	21,955	20,090	18,168	17,535	17,010	16,171	15,507	13,362	12,027	11,030	10,219	9,524	8,909	8,351	**8**
27,877	23,589	21,666	19,679	19,023	18,480	17,608	16,919	14,684	13,288	12,242	11,389	10,656	10,006	9,414	**9**
29,588	25,188	23,209	21,161	20,483	19,922	19,021	18,307	15,987	14,534	13,442	12,549	11,781	11,097	10,473	**10**
31,264	26,757	24,725	22,618	21,920	21,342	20,412	19,675	17,275	15,767	14,631	13,701	12,899	12,184	11,530	**11**
32,909	28,300	26,217	24,054	23,337	22,742	21,785	21,026	18,549	16,989	15,812	14,845	14,011	13,266	12,584	**12**
34,528	29,819	27,688	25,472	24,736	24,125	23,142	22,362	19,812	18,202	16,985	15,984	15,119	14,345	13,636	**13**
36,123	31,319	29,141	26,873	26,119	25,493	24,485	23,685	21,064	19,406	18,151	17,117	16,222	15,421	14,685	**14**
37,697	32,801	30,578	28,259	27,488	26,848	25,816	24,996	22,307	20,603	19,311	18,245	17,322	16,494	15,733	**15**
39,252	34,267	32,000	29,633	28,845	28,191	27,136	26,296	23,542	21,793	20,465	19,369	18,418	17,565	16,780	**16**
40,790	35,718	33,409	30,995	30,191	29,523	28,445	27,587	24,769	22,977	21,615	20,489	19,511	18,633	17,824	**17**
42,312	37,156	34,805	32,346	31,526	30,845	29,745	28,869	25,989	24,155	22,760	21,605	20,601	19,699	18,868	**18**
43,820	38,582	36,191	33,687	32,852	32,158	31,037	30,144	27,204	25,329	23,900	22,718	21,689	20,764	19,910	**19**
45,315	39,997	37,566	35,020	34,170	33,462	32,321	31,410	28,412	26,498	25,038	23,828	22,775	21,826	20,951	**20**
46,797	41,401	38,932	36,343	35,479	34,759	33,597	32,671	29,615	27,662	26,171	24,935	23,858	22,888	21,991	**21**
48,268	42,796	40,289	37,659	36,781	36,049	34,867	33,924	30,813	28,822	27,301	26,039	24,939	23,947	23,031	**22**
49,728	44,181	41,638	38,968	38,076	37,332	36,131	35,172	32,007	29,979	28,429	27,141	26,018	25,006	24,069	**23**
51,179	45,559	42,980	40,270	39,364	38,609	37,389	36,415	33,196	31,132	29,553	28,241	27,096	26,063	25,106	**24**
52,620	46,928	44,314	41,566	40,646	39,880	38,642	37,652	34,382	32,282	30,675	29,339	28,172	27,118	26,143	**25**
54,052	48,290	45,642	42,856	41,923	41,146	39,889	38,885	35,563	33,429	31,795	30,435	29,246	28,173	27,179	**26**
55,476	49,645	46,963	44,140	43,195	42,407	41,132	40,113	36,741	34,574	32,912	31,528	30,319	29,227	28,214	**27**
56,892	50,993	48,278	45,419	44,461	43,662	42,370	41,337	37,916	35,715	34,027	32,620	31,391	30,279	29,249	**28**
58,301	52,336	49,588	46,693	45,722	44,913	43,604	42,557	39,087	36,854	35,139	33,711	32,461	31,331	30,283	**29**

0,001	0,005	0,01	0,02	0,025	0,03	0,04	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	G.L.
59,703	53,672	50,892	47,962	46,979	46,160	44,834	43,773	40,256	37,990	36,250	34,800	33,530	32,382	31,316	**30**
61,098	55,003	52,191	49,226	48,232	47,402	46,059	44,985	41,422	39,124	37,359	35,887	34,598	33,431	32,349	**31**
62,487	56,328	53,486	50,487	49,480	48,641	47,282	46,194	42,585	40,256	38,466	36,973	35,665	34,480	33,381	**32**
63,870	57,648	54,776	51,743	50,725	49,876	48,500	47,400	43,745	41,386	39,572	38,058	36,731	35,529	34,413	**33**
65,247	58,964	56,061	52,995	51,966	51,107	49,716	48,602	44,903	42,514	40,676	39,141	37,795	36,576	35,444	**34**
66,619	60,275	57,342	54,244	53,203	52,335	50,928	49,802	46,059	43,640	41,778	40,223	38,859	37,623	36,475	**35**
73,402	66,766	63,691	60,436	59,342	58,428	56,946	55,758	51,805	49,244	47,269	45,616	44,165	42,848	41,622	**40**
99,607	91,952	88,379	84,580	83,298	82,225	80,482	79,082	74,397	71,341	68,972	66,981	65,227	63,628	62,135	**60**
124,839	116,321	112,329	108,069	106,629	105,422	103,459	101,879	96,578	93,106	90,405	88,130	86,120	84,284	82,566	**80**
137,208	128,299	124,116	119,648	118,136	116,869	114,806	113,145	107,565	103,904	101,054	98,650	96,524	94,581	92,761	**90**
149,449	140,169	135,807	131,142	129,561	128,237	126,079	124,342	118,498	114,659	111,667	109,141	106,906	104,862	102,946	**100**
173,617	163,648	158,950	153,918	152,211	150,780	148,447	146,567	140,233	136,062	132,806	130,055	127,616	125,383	123,289	**120**
197,451	186,847	181,840	176,471	174,648	173,118	170,624	168,613	161,827	157,352	153,854	150,894	148,269	145,863	143,604	**140**

0,45	0,50	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95	0,975	0,98	0,99	0,995	G.L.
0,571	0,455	0,357	0,275	0,206	0,148	0,102	0,064	0,036	0,016	0,004	0,001	0,001	0,000	0,000	**1**
1,597	1,386	1,196	1,022	0,862	0,713	0,575	0,446	0,325	0,211	0,103	0,051	0,040	0,020	0,010	**2**
2,643	2,366	2,109	1,869	1,642	1,424	1,213	1,005	0,798	0,584	0,352	0,216	0,185	0,115	0,072	**3**
3,687	3,357	3,047	2,753	2,470	2,195	1,923	1,649	1,366	1,064	0,711	0,484	0,429	0,297	0,207	**4**
4,728	4,351	3,996	3,655	3,325	3,000	2,675	2,343	1,994	1,610	1,145	0,831	0,752	0,554	0,412	**5**
5,765	5,348	4,952	4,570	4,197	3,828	3,455	3,070	2,661	2,204	1,635	1,237	1,134	0,872	0,676	**6**
6,800	6,346	5,913	5,493	5,082	4,671	4,255	3,822	3,358	2,833	2,167	1,690	1,564	1,239	0,989	**7**
7,833	7,344	6,877	6,423	5,975	5,527	5,071	4,594	4,078	3,490	2,733	2,180	2,032	1,646	1,344	**8**
8,863	8,343	7,843	7,357	6,876	6,393	5,899	5,380	4,817	4,168	3,325	2,700	2,532	2,088	1,735	**9**
9,892	9,342	8,812	8,295	7,783	7,267	6,737	6,179	5,570	4,865	3,940	3,247	3,059	2,558	2,156	**10**
10,920	10,341	9,783	9,237	8,695	8,148	7,584	6,989	6,336	5,578	4,575	3,816	3,609	3,053	2,603	**11**
11,946	11,340	10,755	10,182	9,612	9,034	8,438	7,807	7,114	6,304	5,226	4,404	4,178	3,571	3,074	**12**
12,972	12,340	11,729	11,129	10,532	9,926	9,299	8,634	7,901	7,042	5,892	5,009	4,765	4,107	3,565	**13**
13,996	13,339	12,703	12,078	11,455	10,821	10,165	9,467	8,696	7,790	6,571	5,629	5,368	4,660	4,075	**14**
15,020	14,339	13,679	13,030	12,381	11,721	11,037	10,307	9,499	8,547	7,261	6,262	5,985	5,229	4,601	**15**
16,042	15,338	14,655	13,983	13,310	12,624	11,912	11,152	10,309	9,312	7,962	6,908	6,614	5,812	5,142	**16**

0,45	0,50	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95	0,975	0,98	0,99	0,995	G.L.
17,065	16,338	15,633	14,937	14,241	13,531	12,792	12,002	11,125	10,085	8,672	7,564	7,255	6,408	5,697	**17**
18,086	17,338	16,611	15,893	15,174	14,440	13,675	12,857	11,946	10,865	9,390	8,231	7,906	7,015	6,265	**18**
19,107	18,338	17,589	16,850	16,109	15,352	14,562	13,716	12,773	11,651	10,117	8,907	8,567	7,633	6,844	**19**
20,127	19,337	18,569	17,809	17,046	16,266	15,452	14,578	13,604	12,443	10,851	9,591	9,237	8,260	7,434	**20**
21,147	20,337	19,548	18,768	17,984	17,182	16,344	15,445	14,439	13,240	11,591	10,283	9,915	8,897	8,034	**21**
22,166	21,337	20,529	19,729	18,924	18,101	17,240	16,314	15,279	14,041	12,338	10,982	10,600	9,542	8,643	**22**
23,185	22,337	21,510	20,690	19,866	19,021	18,137	17,187	16,122	14,848	13,091	11,689	11,293	10,196	9,260	**23**
24,204	23,337	22,491	21,652	20,808	19,943	19,037	18,062	16,969	15,659	13,848	12,401	11,992	10,856	9,886	**24**
25,222	24,337	23,472	22,616	21,752	20,867	19,939	18,940	17,818	16,473	14,611	13,120	12,697	11,524	10,520	**25**
26,240	25,336	24,454	23,579	22,697	21,792	20,843	19,820	18,671	17,292	15,379	13,844	13,409	12,198	11,160	**26**
27,257	26,336	25,437	24,544	23,644	22,719	21,749	20,703	19,527	18,114	16,151	14,573	14,125	12,879	11,808	**27**
28,274	27,336	26,419	25,509	24,591	23,647	22,657	21,588	20,386	18,939	16,928	15,308	14,847	13,565	12,461	**28**
29,291	28,336	27,402	26,475	25,539	24,577	23,567	22,475	21,247	19,768	17,708	16,047	15,574	14,256	13,121	**29**
30,307	29,336	28,386	27,442	26,488	25,508	24,478	23,364	22,110	20,599	18,493	16,791	16,306	14,953	13,787	**30**
31,323	30,336	29,369	28,409	27,438	26,440	25,390	24,255	22,976	21,434	19,281	17,539	17,042	15,655	14,458	**31**
32,339	31,336	30,353	29,376	28,389	27,373	26,304	25,148	23,844	22,271	20,072	18,291	17,783	16,362	15,134	**32**
33,355	32,336	31,337	30,344	29,340	28,307	27,219	26,042	24,714	23,110	20,867	19,047	18,527	17,074	15,815	**33**
34,371	33,336	32,322	31,313	30,293	29,242	28,136	26,938	25,586	23,952	21,664	19,806	19,275	17,789	16,501	**34**
35,386	34,336	33,306	32,282	31,246	30,178	29,054	27,836	26,460	24,797	22,465	20,569	20,027	18,509	17,192	**35**
40,459	39,335	38,233	37,134	36,021	34,872	33,660	32,345	30,856	29,051	26,509	24,433	23,838	22,164	20,707	**40**
60,713	59,335	57,978	56,620	55,239	53,809	52,294	50,641	48,759	46,459	43,188	40,482	39,699	37,485	35,534	**60**
80,927	79,334	77,763	76,188	74,583	72,915	71,145	69,207	66,994	64,278	60,391	57,153	56,213	53,540	51,172	**80**
91,023	89,334	87,666	85,993	84,285	82,511	80,625	78,558	76,195	73,291	69,126	65,647	64,635	61,754	59,196	**90**
101,115	99,334	97,574	95,808	94,005	92,129	90,133	87,945	85,441	82,358	77,929	74,222	73,142	70,065	67,328	**100**
121,285	119,334	117,404	115,465	113,483	111,419	109,220	106,806	104,037	100,624	95,705	91,573	90,367	86,923	83,852	**120**
141,441	139,334	137,248	135,149	133,003	130,766	128,380	125,758	122,748	119,029	113,659	109,137	107,815	104,034	100,655	**140**

7.5 Tabla V

Variables Aleatorias Discretas (Lind, 2012).

Variable discreta	Fórmula	μ	Σ²	Evento
Uniforme discreta	$F(x) = \frac{1}{x}$	$E(x) = \frac{\sum_{i=1}^{k} xi\, f(x)}{k}$	$\sigma^2 = E(x)^2 - [E(x)]^2$	$x = x_1, x_2, x_3 \ldots\ldots\ldots x_n$
Bernoulli	$F(x) = p * q^{1-p}$ *p* probabilidad de éxito *(1 – p)* probabilidad de fracaso	$\mu = p$	$\sigma^2 = p * q$	$pe^t + q$
Binomial	$B(x) = \binom{n}{x} p^X q^{N-X}$	$\mu = n * p$	$\sigma^2 = n * p * q$	$(pe^t + q)^n$
Binomial negativa	$B^*(x) = \binom{x-1}{k-1} p^k q^{x-k}$ $k = k, k+1, k+2 \ldots..$	$\mu = \frac{k}{p}$	$\sigma^2 = \frac{k(1-p)}{p^2}$	$x = \left(\frac{pe^t}{1-(1-p)e^t}\right)^t$
Multinomial	$f(x) = \binom{\boldsymbol{n}}{\boldsymbol{x_1}, \boldsymbol{x_2} \ldots . \boldsymbol{x_k}} \boldsymbol{p_1}^{\boldsymbol{x1}} * \boldsymbol{p_2}^{\boldsymbol{x2}} \boldsymbol{p_k}^{\boldsymbol{xk}}$	$\mu = n * p_k$	$\mu = n * p_k(1 - p_k)$	No se da.
Geométrica	$G(x) = p * q^{x-1}$ $x = 1, 2, 3, \ldots n$	$\mu = \frac{1}{p}$	$\sigma^2 = \frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Hipergeométrica	$h(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$	$\mu = \frac{n * k}{N}$	$\sigma^2 = \frac{k(N-k) * n(N-n)}{N^2(N-1)}$	No se da.
Poisson	$P(\boldsymbol{x};\, \boldsymbol{\lambda t}) = \left(\frac{\boldsymbol{e^{-\lambda t}(\lambda t)^x}}{\boldsymbol{x}!}\right)$ $x = 1, 2, 3, .. n$	$\mu = \boldsymbol{\lambda}$	$\sigma^2 = \boldsymbol{\lambda}$	$\boldsymbol{e^{\lambda(e^t - 1)}}$

Variables Aleatorias Contínuas

Variable contínua	Fórmula	M	S^2	Evento
Uniforme contínua	$F(x) = \begin{cases} \dfrac{1}{b-a} \end{cases}$ $a \leq x \leq b$	$M = \dfrac{a+b}{2}$	$S^2 = \dfrac{(b-a)^2}{12}$	$\dfrac{e^{tb} - e^{ta}}{t(b-a)}$
Gamma	$F(x) = \dfrac{x^{\propto -1} e^{-x/\beta}}{\beta^{\propto} \Gamma(\propto)}$ $0 \leq x \leq \infty$	$M = \propto * \boldsymbol{\beta}$	$S^2 = \propto * \beta^2$	$\dfrac{1}{(1-\beta t)^{\propto}}$
Exponencial	$F(x) = \dfrac{1}{\beta} e^{-x/\beta}$ $0 \leq x \leq \infty$	$M = \boldsymbol{\beta}$	$S^2 = \beta^2$	$\dfrac{1}{(1-\beta t)}$
Normal o Gaussiana	$n(x) = \dfrac{(e^{-\frac{1}{2\sigma^2} * (x-\mu)^2})}{\sqrt{2\pi\sigma}}$ $-\infty \leq x \leq \infty$	μ	σ^2	$e^{\mu t + (\frac{t2\sigma2}{2})}$
T student	$\beta(t) = \dfrac{\Gamma\left[\frac{(v+1)}{2}\right]}{\Gamma\left(\frac{v}{2}\right)\sqrt{u}} \left(1 + \dfrac{t^2}{v}\right)^{-\frac{v+1}{2}}$	$M = \dfrac{\propto}{\propto + \beta}$	$M = \dfrac{\propto \beta}{(\propto + \beta)^2 (\propto + \beta + 1)}$	No se da.
Chi- cuadrada (x^2)	$F(x) = \dfrac{x^{n/2-1} e^{-x/2}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})}$ $0 \leq x^2 \leq \infty$	$M = \boldsymbol{n}$	$S^2 = 2n$	$(1-2t)^{-n/2}$

7.6 Integral definida o áreas bajo la curva normal

En la geometría elemental se entiende mediante las fórmulas para hallar el área de cualquier región limitada por una poligonal cerrada. Ahora bien, si una área o región está delimitada por alguna línea curva, como es el círculo, el área se expresa como un límite de las áreas de poligonales "próximas" (Martínez, 2012).

Dada una función y = f (x) integrable en un intervalo [a,b], el área de la región limitada por la función, el eje de las ordenada y las rectas x = a y x = b se define como:

$$A = \int_a^b |f(x)|\, dy$$

Ejemplo:

Calcular el área de la región limitada por la gráfica de la siguiente función f y el eje x en el siguiente intervalo indicado:

a. $f(x) = x\,(\ln x)^2$;en [1,e]

b. $f(x) = c^{-x}\,|\text{sen } x|$; en [0,2π]

Donde:

a. $f(x) = x\,(\ln x)^2$;en [1,e]

La función expresada con la integral.

$$A = \int_1^e x\,(\ln x)^2\, dx$$

Análisis de la antiderivada.

$$\left[\frac{x^2}{2} * \ln(x)^2 - \frac{x^2}{2} * \ln(x)^2 + \frac{x^2}{4}\right]\begin{matrix} e \\ 1 \end{matrix}$$

Reemplazo de la variable x.

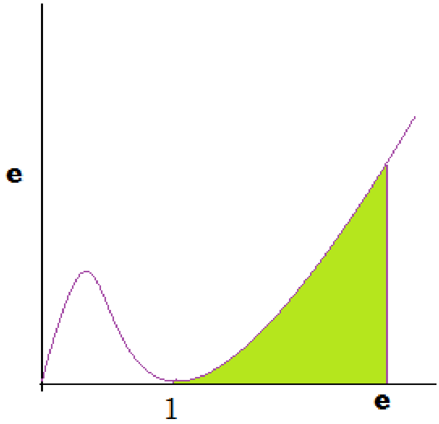
Ln(1)= 0

Ln(e)=1

$$\left(\frac{e^2}{2} * \ln(e)^2 - \frac{e^2}{2} * \ln(e)^2 + \frac{e^2}{4}\right) - \left(\frac{e^2}{2} * \ln(1)^2 - \frac{e^2}{2} * \ln(1)^2 + \frac{1^2}{4}\right)$$

$$\frac{e^2}{4} - \frac{1}{4} = \frac{e^2 - 1}{4}$$

Representación gráfica de la expresión.



b. f(x) = e^{-x} |sen x|; en [0,2π]

La función expresada con la integral.

$$A = \int_0^{2\pi} \frac{e^{-x}}{2} |sen\ x| dx$$

Demarcacion de los límites.

$$\int_0^{\pi} e^{-x}\, sen\ x\ dx + \int_0^{2\pi} -e^{-x}\, sen\ x\ dx$$

Análisis de la antiderivada.

$$\left[\frac{e^{-x}}{2}(\sin x + \cos x)\right]_0^{\pi} + \left[\frac{e^{-x}}{2}(\sin x + \cos x)\right]_{\pi}^{2\pi}$$

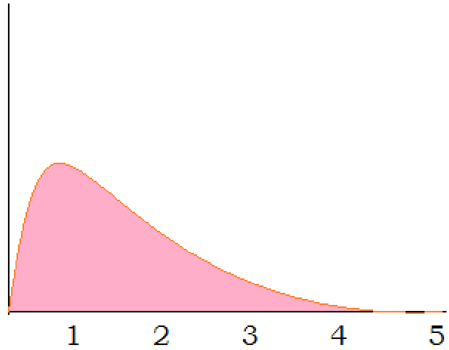
Reemplazo de la variable x.

$$\left[\frac{e^{-\pi}}{2}(\sin \pi + \cos \pi)\right] - \left[\frac{e^{-0}}{2}(\sin 0 + \cos 0)\right] +$$
$$\left[\frac{e^{-2\pi}}{2}(\sin 2\pi + \cos 2\pi)\right] - \left[\frac{e^{-\pi}}{2}(\sin \pi + \cos \pi)\right]$$

- $\sin(0) = 0$
- $\cos(0) = 1$
- $\sin \pi = 0$
- $\cos \pi = -1$

$$\frac{(e^{-\pi}+1)^2}{2}$$

Representación gráfica de la expresión.



Para profundizar el proceso de integración vease también Cálculo Aplicado para Administración Economía y Negocios (Laurance Hoffmann, 2006).

Áreas bajo la curva normal
No importa cuáles sean los valores de la curva, para una distribución de probabilidad normal, el área total bajo la curva es 1.00, de manera que podemos pensar en áreas bajo la curva como si fueran probabilidades.

Matemáticamente es verdad que (Anderson, 2012):

1.Aproximadamente 68% de todos los valores de una población normalmente distribuida se encuentra dentro de desviación estándar de la media.

2. Aproximadamente 95.5 % de todos los valores de una población normalmente distribuida se encuentra dentro de desviación estándar de la media.

3. Aproximadamente 99.7 % de todos los valores de una población normalmente se encuentra distribuida dentro de desviación estándar de la media.

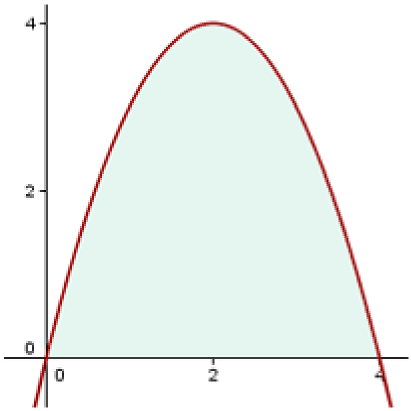
También podemos encontrar la tabla que indica el área bajo la curva normal estándar que corresponde a $P(Z < z)$ para valores de z que van de –3.49 a 3.49.

Ejemplo:

Calcular el área de la curva que se encuentra limitado por:
$y = 4x - x^2$

En primer lugar los puntos de barrido de la gráfica con el eje cero para incorporar la curva:

$y = 4x - x^2$
$y = x\,(4 - x)$ factor común
$x = 0\ ;\ x = 4$



Se calcula la integral definida de 0 a 4:

$$A = \int_0^4 (4x - x^2)\,dx$$

La expresión integrada se expresa como:

$$\left[\frac{4x^2}{2} - \frac{x^3}{3}\right] = \left[2x^2 - \frac{x^3}{3}\right]$$

Suma algebraica de la expresión y remplazo de valores en la variable x.

$$\frac{6x^2 - x^3}{3} = \frac{6(4)^2 - (4)^3}{3} - 0 = \frac{32}{3}$$

7.7 Diccionario de términos estadísticos

Términos estadísticos de uso común en las unidades de estudio (Instituto Nacional de Estadísticas e Informática , 2006).

A

AMPLITUD DE INTERVALO.- Destacado también como amplitud de clase, es la diferencia entre dos cotas o límites de un intervalo.

ANÁLISIS DE VARIANZA.- Es un método para contrastar dos o más medias de n grupos examinando la varianza de los datos, tanto entre n grupos como dentro de ellos.

ASIMETRÍA.- Es la inexactitud de simetría entre los datos de una distribución.

B

BASE DEL ÍNDICE.- Es la dimensión manipulada como unidad de referencia, contra la cual se hacen todas las asimilaciones de la variable en estudio.

C

CARTOGRAMAS.- Es un tipo de gráfico donde se muestra datos estadísticos sobre una base geográfica como mapas.

CENSO.- Es una averiguación estadística que consiste en el cálculo de la totalidad de los elementos que componen la población por investigar.

CICLO.- Variaciones o fluctuaciones cíclicas.

CLASE MEDIANA.- Es la clase o intervalo al que pertenece el valor de la mediana.

CLASE MODAL.- Es la clase o intervalo que tiene la mayor frecuencia.

CLASE O CATEGORÍA.- Se designa así a la característica o los intervalos construidos favorablemente para agrupar la información.

CODIFICACIÓN.- Es asignar números o claves a la información para preparar el procesamiento de datos.

COEFICIENTE DE VARIACIÓN.- Es una medida de dispersión y se calcula dividiendo la desviación típica entre la media aritmética. Se explica como porcentaje, por lo que nos autorizará decidir entre dos muestras, cuál es la que presenta mayor dispersión.

COMBINACIONES.- Radica en seleccionar diferentes agrupaciones de r elementos de un total de n objetos sin tomar en cuenta el orden, y el número de combinaciones.

CONGLOMERADO.- Es una subpoblación o agrupación que reúne características presentes en la población.

CUARTIL.- Es una medida de posición no central. Los cuartiles son los tres valores que fragmentan la distribución en cuatro partes asimiles, es decir, dentro de cada cual están incluidos el 25% de los datos de la distribución:

Q 1 Representa el primer cuartil es el 25% de la distribución.

Q 2 Representa el segundo cuartil y es el 50% de la distribución y este valor es igual a la mediana.

Q 3 Representa el tercer cuartil y es el 75% de la distribución.

CUESTIONARIO.- Es uno de los instrumentos más utilizado para recolectar datos.

CURTOSIS.- Es una medida de forma. Se concretan 3 tipos de distribuciones según su grado de curtosis: Distribución mesocúrtica, distribución leptocúrtica y distribución platicúrtica.

CURVA NORMAL.- También llamada curva o campana de Gauss, es una distribución simétrica de mediciones, con el mismo valor de casos a distancias determinadas tanto por debajo como por encima de la media.

D

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

DATO CUALITATIVO.- Es aquel que significa alguna característica de los elementos de una muestra o una población son datos NO NUMÉRICOS.

DATO CUANTITATIVO.- Es aquel dato numérico que simboliza aspectos de una muestra o una población que es medible o que se puede contar.

DESVIACIÓN ESTÁNDAR.- Conocida también como desviación típica, es una medida de dispersión que se logra mediante la raíz cuadrada de la varianza.

DIAGRAMA.- Es un dibujo o representación gráfica que se utiliza para representar un objeto, demostrar la relación entre elementos o el valor de una magnitud.

DIAGRAMA DE BARRAS.- Es un gráfico utilizado para representar la distribución de frecuencias de una variable cualitativa y cuantitativa discreta.

DIAGRAMA DE SECTORES.- Es un gráfico usado para personificar la distribución de frecuencias relativas de una variable cualitativa.

E

ENCUESTA.- Es un método de recolección de datos.

ENTREVISTA.- Es un método de recolección de datos. Consiste en una serie de preguntas realizadas.

ESPACIO MUESTRAL.- Es el conjunto de todos los resultados posibles de un experimento aleatorio.

ESTADÍSTICA.- Es la ciencia que abarca una serie de métodos y procedimientos asignados a la recopilación, tabulación, procesamiento, análisis e interpretación de datos cuantitativos y cualitativos. Se divide en dos tipos: Estadística descriptiva y Estadística inferencial.

ESTADÍSTICA DESCRIPTIVA.- Rama de la ciencia estadística que se encarga desde la recopilación, procesamiento y análisis de la información siendo sus conclusiones válidas sólo para el grupo analizado.

ESTADÍSTICA INFERENCIAL.- Rama de la ciencia estadística que provee métodos y operaciones que permiten obtener conclusiones para una población a partir del estudio de una o más muestras representativas.

ESTIMADOR.- Es un estadístico utilizado para estimar un parámetro.

EXPERIMENTO.- Es un método de exploración mediante el cual se establece la incidencia de variables independientes sobre la variable dependiente.

EXPERIMENTO ALEATORIO.- Es el hecho que implica la observación de los valores de una variable aleatoria, donde puede dar varios resultados, sin que pueda ser previsible expresar con certeza cuál de éstos va a ser observado en la ejecución del experimento.

F

FUENTES DE DATOS.- Medios de donde proviene la información y se reúnen de diferentes fuentes de información ya existente.

G

GRADOS DE LIBERTAD.- Es un estadístico calculado en base a n datos, se refiere al número de cantidades independientes, menos el número de restricciones que se asocian a las observaciones y el estadístico. Simbólicamente se representa por gl.

H

HISTOGRAMA.- Gráfico usado para personificar la distribución de frecuencias de una variable continúa. Describe el procedimiento de un grupo de datos en todo lo relacionado a su tendencia central, forma y dispersión.

I

INDEPENDENCIA ESTADÍSTICA.- Se entiende que dos variables X e Y son independientes, estadísticamente, cuando la frecuencia relativa conjunta es igual al producto de las frecuencias relativas marginales y si esta condición no se efectúa para todos los valores, se dice que hay dependencia estadística.

ÍNDICE.- Es la relación formulada en porcentaje entre el precio, cantidad o valor de un bien y servicio o grupo de bienes y servicios, en una etapa de estudio y el precio, cantidad o valor del mismo bien y servicio o grupo de bienes y servicios en la etapa de referencia o período base.

L

LÍMITE INFERIOR.- Es el menor valor de un intervalo de clase.

LÍMITE SUPERIOR.- Es el mayor valor de un intervalo de clase.

M

MUESTREO.- Es una agrupación de métodos y operaciones estadísticos asignados a la selección de una o más muestras es la técnica seguida para elegir muestras.

N

NÚMERO ÍNDICE.- Es aquella medida estadística que consiente estudiar las variaciones que se producen en una magnitud simple o compleja con relación al tiempo o al espacio es decir, se va a comparar dos escenarios, una de las cuales se considera de referencia, llamado también período base.

O

ORDEN DE DATOS *(véase como fuente de datos)*

P

PARÁMETRO.- Es cualquier valor exclusivo de la población.

PERMUTACIONES.- Son varias disposiciones de los elementos en que se pueden ordenar los objetos. El número de permutaciones de n objetos se obtiene como la factorial de n!

PERCENTIL.- Es el valor que se origina al dividir el conjunto de datos en 100 partes iguales. Cada parte representa al 1% del total.

PICTOGRAMAS.- Son gráficos vistosos, equivalentes a los gráficos de barras, pero empleando un dibujo alusivo al tema que representa.

POBLACIÓN FINITA.- Es aquella en la que es viable enumerar físicamente los elementos que pertenecen a la población.

POBLACIÓN INFINITA.- Es aquella en la que no es viable enumerar físicamente los elementos que pertenecen a la población.

POBLACIÓN O UNIVERSO.- Es cualquier conjunto de unidades o elementos claramente definidos, en el espacio y el tiempo.

PROBABILIDAD.- Es un número que se le fija a un suceso como una medida de su incertidumbre.

PROMEDIO SIMPLE.- Es cualquier medida de posición de tendencia central. Cuando se consigue sumando los datos y dividendo entre el número de ellos.

PROMEDIO PONDERADO.- Es un número conocido también como media aritmética ponderada.

Q

QUINTIL.- Es un fractil se logran dividiendo al conjunto de datos en cinco partes iguales cada parte representa el 20% del total. Se pueden calcular 4 quintiles.

R

RANGO.- Conocido como recorrido, es un número que calcula la amplitud de los valores de un conjunto de datos y se desarrolla por medio de la diferencia entre el valor mayor y el valor menor.

S

SESGO.- Se denomina así a la asimetría que presenta una distribución de frecuencias.

SUCESO CIERTO.- Se conoce también como suceso seguro.

SUCESO IMPOSIBLE.- Es aquel evento que no se realiza nunca. Se designa por un Ø.

T

TABLA DE DATOS AGRUPADOS.- Es un ajuste matricial que contiene el número de veces (frecuencia) que aparece un dato, de acuerdo a las clases de interés detalladas (puede ser intervalos).

TASA.- Es la relación del número de casos, frecuencias o eventos de una categoría entre el número total de observaciones, multiplicada por un múltiplo de 10.

U

UNIDAD DE MUESTREO.- Es la unidad estadística que se prefiere para constituir la muestra.

V

VARIABLE.- Es una característica de la población o de la muestra cuya medida puede cambiar de valor.

VARIABLE ALEATORIA.- Conocida igualmente como variable estocástica o probabilística.

VARIABLE CONTÍNUA.- Es una variable cuantitativa.

VARIABLE CUALITATIVA.- Es aquella que simboliza cualidades, atributos o características no numéricas.

VARIABLE CUANTITATIVA.- Es aquella característica que es posible representar numéricamente. Éstas pueden ser contínuas y discretas.

VARIABLE DISCRETA.- Es una variable cuantitativa. Es la característica de la población, cuyos valores están personificados por medio del conjunto de los números naturales.

VARIANZA.- Es una medida de dispersión de la información que se obtiene como el promedio de los cuadrados de las desviaciones de los valores de la variable en relación a su media aritmética.

7.8 ÍNDICE DE NOMBRES

Bibliografía

Anderson, S. W. (2012). *Estadítica para negocios y economía*. Cincinnati, Estados Unidos: Cengage Learning.

IBM Corp. (junio de 2017). *IBM Marketplace*. Recuperado de https://www.ibm.com/us-en/marketplace/spss-statistics

Instituto Nacional de Estadísticas e Informática (2006). *Glosario básico de terminos estadísticos*. Lima, Perú: INEI.

Instituto Nacional de Estadísticas y Censos (INEC). (22 de Mayo de 2015). *Ecuador en cifras*. Recuperado de http://www.ecuadorencifras.gob.ec

Laurance Hoffmann, G. B. (2006). *Cálculo aplicado para administración economía y negocios.* México, México: Mc-Grawhill.

Levin, R. B. (2010). *Estadística para administración y economía.* New York, Estados Unidos: Pearson.

Lind, M. W. (2012). *Estadística aplicada a los negocios y la economía.* Michigan, Estados Unidos: Mc Graw Hill.

Martínez, C. (2012). *Estadística y muestreo.* Medellín, Colombia: Ecoe Ediciones.

Math Works (Agosto de 2017). *Matlab y simulink*. Recuperado de https://es.mathworks.com/

NC State University (julio de 2017). *Statistical analysis system*. Recuperado de Universidad Estatal de Carolina del Norte: https://www.sas.com/es_ec/company-information.html

Ruíz Muñoz, David, Universidad Pablo de Olavide (Febrero de 2008). *Enciclopedia Virtual Eumed*. Recuperado de http://www.eumed.net/cursecon/libreria/drm/0.htm

Statsoft(Agosto de 2017). *Statistica Features*. Recuperado de http://www.statsoft.com/Products/Statistica-Features

Walpole, R. H. (1999). *Probabilidad y estadística para ingenieros.* México, México: Pearson Education.

Webster Allen (2008). *Estadística aplicada a los negocios y la economía.* New York, Estados Unidos: Mc Graw Hill.

El texto ha sido diseñado para ser utilizado a nivel de pequeños, medianos y grandes empresarios, así como también por estudiantes de licenciatura de la carrera de Economía de las diferentes Universidades del país, manejando un criterio sencillo, dinámico y espontáneo. Con un esquema ordenado explica cómo la Estadística permite utilizar herramientas de análisis de datos para la toma de decisiones, donde se abarcan conceptos de Agrupación y Presentación de los Datos Estadísticos, Análisis de las Medidas de Tendencia Central y Dispersión, Números índices, Teoría de Conjuntos Cordinatoria y de Probabilidades, Variables Aleatorias y Distribución Conjunta.

ISBN: 978-9942-759-52-8
9789942759528

CIDE
EDITORIAL
Cod. 9942-8632