ESTADÍSTICA Y PROBABILIDADES

© NERY ELISABETH GARCÍA PAREDES
ALEXANDER FERNANDO HARO SARANGO
LIZETH FERNANDA SILVA GODOY
FREDIN FERNANDO POZO PARRA
STALIN GABRIEL SALGUERO GUALPA

CASA EDITORA DEL POLO

“ESTADÍSTICA Y PROBABILIDADES”

Nery Elisabeth García - Paredes
Alexander Fernando Haro – Sarango
Lizeth Fernanda Silva - Godoy
Fredin Fernando Pozo - Parra
Stalin Gabriel Salguero - Gualpa

CASA EDITORA DEL POLO

ESTADÍSTICA Y PROBABILIDADES

Nery Elisabeth García – Paredes. Docente en la carrera de Biotecnología de la Universidad Técnica de Cotopaxi; y docente la carrera de Administración de Empresas y Negocios Internacionales de la Pontificia Universidad Católica del Ecuador sede Ambato.

Alexander Fernando Haro Sarango. Docente de Administración Financiera y como Coordinador de la carrera Tecnológica Superior Universitaria en Administración de Empresas e Inteligencia de Negocios en el Instituto Superior Tecnológico España (ISTE), Ambato – Ecuador.

Lizeth Fernanda Silva – Godoy. Docente de Ciencias Exactas en Cálculo vectorial y Álgebra Lineal en los departamentos de Energía y Mecánica, Eléctrica y Electrónica, Ciencias de la Computación (ESPE-L), Latacunga -Ecuador.

Fredin Fernando Pozo – Parra. Docente de Algebra lineal y Cálculo Diferencial e Integral, de la Facultad de Ciencias Agrarias y Forestales (UTEQ), Quevedo – Ecuador.

Stalin Gabriel Salguero – Gualpa. Asesor académico, capacitador en la institución privada ESCAM.

CASA EDITORA DEL POLO

Casa Editora del Polo - CASEDELPO CIA. LTDA.
Departamento de Edición

Editado y distribuido por:

Editorial: Casa Editora del Polo
Sello Editorial: 978-9942-816
Manta, Manabí, Ecuador. 2019
Teléfono: (05) 6051775 / 0991871420
Web: www.casedelpo.com
ISBN: 978-9942-621-59-7
DOI: https://doi.org/10.23857/978-9942-621-59-7

Revisión, Ortografía y Redacción:
Lic. Jessica Mero Vélez

Diseño de Portada:
Michael Josué Suárez-Espinar

Diagramación:
Ing. Edwin Alejandro Delgado-Veliz

Director Editorial:
Dra. Tibisay Milene Lamus-García

Constancia de Arbitraje

La Casa Editora del Polo, hace constar que este libro proviene de una investigación realizada por los autores, siendo sometido a un arbitraje bajo el sistema de doble ciego (peer review), de contenido y forma por jurados especialistas. Además, se realizó una revisión del enfoque, paradigma y método investigativo; desde la matriz epistémica asumida por los autores, aplicándose las normas APA, Sexta Edición, proceso de anti plagio en línea Plagiarisma, garantizándose así la cientificidad de la obra.

Comité Editorial

Contenido

PROLOGO

Bienvenidos a un viaje fascinante y enriquecedor hacia el mundo de la estadística y la teoría de probabilidades. En las páginas que siguen, nos embarcaremos en un recorrido por los misteriosos y apasionantes territorios de los números, las incertidumbres y las predicciones. A través de este libro, desentrañaremos los conceptos fundamentales que subyacen a estas disciplinas y descubriremos cómo han moldeado nuestro entendimiento del mundo y han revolucionado una multitud de campos, desde la ciencia hasta la economía y más allá.

La estadística y la teoría de probabilidades son como las dos caras de una moneda; están intrínsecamente entrelazadas y se complementan mutuamente para proporcionar una comprensión profunda de los fenómenos aleatorios y la variabilidad en nuestros datos. En este libro, exploraremos cómo estas herramientas han evolucionado a lo largo de los siglos, desde sus raíces en los juegos de azar y la astronomía hasta su centralidad en la toma de decisiones informadas en la era moderna.

Nuestro viaje comenzará con los conceptos básicos. Descubriremos cómo las probabilidades, inicialmente vistas como meros números en juegos de dados, se convirtieron en la base de la inferencia estadística, permitiéndonos hacer predicciones y tomar decisiones en un mundo donde la certeza es rara vez alcanzable. A medida que profundicemos en la teoría de probabilidades, desentrañaremos los misterios detrás de distribuciones

como la binomial, la normal y la exponencial, y cómo estas formas bellamente abstractas describen la realidad con sorprendente precisión.

La estadística, por su parte, nos brindará las herramientas necesarias para extraer significado de los datos del mundo real. Aprenderemos a resumir y visualizar conjuntos de información aparentemente caóticos, revelando patrones y tendencias que de otro modo podrían haber pasado desapercibidos. Mediante la inferencia estadística, exploraremos cómo extraer conclusiones sobre poblaciones enteras a partir de muestras limitadas, y cómo evaluar la confiabilidad de nuestras estimaciones.

Pero la estadística no se trata solo de números y fórmulas. También implica un elemento crítico de pensamiento crítico. En este libro, aprenderemos a ser cautelosos ante las trampas comunes, como la correlación incorrectamente interpretada y las causalidades erróneas. A medida que desarrollamos nuestras habilidades estadísticas, también cultivaremos un escepticismo saludable hacia las afirmaciones basadas en datos y una apreciación por el valor de diseñar experimentos rigurosos.

En las páginas que siguen, encontrará ejemplos del mundo real que ilustran cómo estas herramientas pueden aplicarse en situaciones tan diversas como la medicina, la economía, la biología y más. Aprenderemos

a enfrentar problemas del mundo real, desde la predicción del clima hasta el análisis de encuestas de opinión pública, utilizando la estadística y la teoría de probabilidades como nuestras guías confiables en un terreno a menudo turbulento.

Así que los invito a sumergirse en las profundidades de la estadística y la teoría de probabilidades. Ya sea que sea un estudiante ansioso por aprender, un profesional curioso o simplemente alguien que busca comprender mejor el funcionamiento del mundo que lo rodea, este libro le proporcionará las bases para navegar por un océano de datos con confianza y sabiduría.

¡Prepárese para un viaje emocionante mientras desentrañamos los secretos de la estadística y la teoría de probabilidades juntos!

INTRODUCCIÓN

En un mundo caracterizado por su complejidad y variabilidad, la necesidad de comprender y abordar la incertidumbre se vuelve fundamental. Desde la toma de decisiones empresariales hasta la investigación científica y la planificación gubernamental, enfrentamos constantemente situaciones en las que los datos parecen enigmáticos y las predicciones son esquivas. Es aquí donde la estadística y la teoría de probabilidades entran en juego como las herramientas más poderosas para iluminar las sombras de la incertidumbre y guiar nuestros pasos hacia el terreno de la comprensión y la toma de decisiones informadas.

La estadística y la teoría de probabilidades son mucho más que conjuntos de fórmulas y gráficos. Son los cimientos sobre los cuales se erige el edificio del conocimiento empírico. Desde los días de los filósofos griegos hasta la era de la inteligencia artificial, estas disciplinas han sido los faros que han guiado a la humanidad a través de la bruma de lo desconocido. En este libro, nos embarcaremos en un viaje a través de sus historias, principios y aplicaciones, y desentrañaremos cómo han transformado tanto nuestra comprensión del mundo como la manera en que tomamos decisiones informadas en la vida cotidiana.

La estadística es la ciencia que nos permite sintetizar, analizar y entender conjuntos de datos complejos. A través de técnicas como la descripción de datos, la inferencia estadística y el diseño de experimentos, la estadística

arroja luz sobre patrones ocultos, revela relaciones causales y proporciona una base sólida para la toma de decisiones fundamentadas. Desde la investigación médica que evalúa la eficacia de un nuevo tratamiento hasta la optimización de procesos industriales, la estadística es la herramienta que nos permite convertir montañas de información en conocimiento utilizable.

La teoría de probabilidades, por otro lado, es la disciplina que explora los conceptos matemáticos detrás de la incertidumbre y el azar. Desde las probabilidades simples de lanzar una moneda hasta los modelos complejos de sistemas caóticos, la teoría de probabilidades nos proporciona las herramientas para cuantificar y entender las incertidumbres inherentes a eventos aleatorios. Estas probabilidades se entrelazan con la estadística, permitiéndonos hacer inferencias sobre poblaciones y muestras, y fundamentando nuestra comprensión de los fenómenos estocásticos en el mundo.

En este libro, no solo exploraremos los conceptos teóricos y las técnicas prácticas de la estadística y la teoría de probabilidades, sino que también examinaremos cómo han impactado en una multitud de campos. Desde la genética hasta la economía, desde la ciencia política hasta el análisis financiero, estas disciplinas se entrelazan con nuestra vida cotidiana de formas sorprendentes. A través de ejemplos reales y casos de estudio, veremos cómo estas herramientas han resuelto problemas complejos y han impulsado avances en todas

las áreas del conocimiento.

Al sumergirse en las páginas que siguen, los lectores serán llevados de la mano a través de un viaje que explora la esencia misma de la incertidumbre y cómo la estadística y la teoría de probabilidades han iluminado su camino. Ya sea que sea un novato en busca de una introducción sólida o un experto ansioso por descubrir nuevos enfoques, este libro tiene algo que ofrecer a todos los ávidos de conocimiento.

CAPÍTULO I

ESTADÍSTICA

La estadística consiste en métodos, procedimientos y fórmulas que permiten recopilar información para su posterior análisis y extraer de ella las conclusiones adecuadas. Podemos decir que esto es ciencia de datos, y su objetivo principal es mejorar la comprensión de los hechos a partir de la información disponible.

El origen de la palabra "estadística" suele atribuirse al economista Gottfried Achenwall (Prusia, 1719-1772), quien entendía la estadística como "la ciencia de las cosas pertenecientes al Estado".

Debe tener en cuenta que la estadística NO es una rama de las matemáticas. Utiliza las herramientas de las matemáticas de la misma manera que la física, la ingeniería o la economía, pero eso no las convierte en parte de las matemáticas. Es cierto que están muy relacionadas, pero la estadística y las matemáticas son disciplinas diferentes.

La estadística es una disciplina que estudia el comportamiento de un conjunto de datos en un contexto determinado utilizando una serie de métodos aritméticos. Proporciona información basada en el análisis de un grupo de datos relativamente pequeño que refleja la naturaleza del grupo más grande.

La estadística es una de las áreas más importantes de cualquier campo de estudio, desde la ciencia hasta los negocios, ya que brinda seguridad y precisión en el análisis de los datos de investigación. Por eso, en este

artículo explicaremos qué es, cuál es su finalidad, cómo aplicarlo correctamente y daremos algunos ejemplos instructivos.

1.1 Introducción

La estadística es un conjunto de métodos, normas y reglas que se utilizan para estudiar el comportamiento de ciertos grupos de datos, de acuerdo con un enfoque de investigación, para que se puedan tomar decisiones y sacar conclusiones sobre una situación, un fenómeno o, en un contexto empresarial, las oportunidades en el entorno o la aceptación y demanda de un producto o servicio.

Los métodos estadísticos se refieren a la recopilación, organización y análisis de la información, y su posterior adaptación o representación a través de tablas y gráficos, de forma sistemática, para facilitar la interpretación de los datos, así como permitir la comparación entre ellos y la obtención de datos relevantes. conclusión investigativa.

El concepto de estadística se extiende a la descripción de la actividad de unos datos ya la deducción y evaluación de conclusiones, a partir de los resultados que arrojan las muestras y las operaciones matemáticas que se les aplican.

1.1.1¿Para qué sirve la estadística?

La estadística, en su esencia, es mucho más que un conjunto de números y gráficos. Es una herramienta

poderosa que nos brinda la capacidad de comprender y analizar el mundo a nuestro alrededor de una manera profunda y significativa. A través de sus técnicas y métodos, la estadística se convierte en el lente a través del cual podemos explorar patrones, tomar decisiones informadas y descubrir verdades ocultas en los datos.

En primer lugar, la estadística nos permite organizar y resumir grandes cantidades de información de manera coherente y concisa. Imagina tener un conjunto de datos que representan, por ejemplo, las ventas de una empresa a lo largo de varios años. Sin la estadística, estos números podrían parecer abrumadores e incomprensibles. Sin embargo, mediante la aplicación de técnicas de resumen, como el cálculo de promedios, medianas y desviaciones estándar, podemos obtener una imagen clara de las tendencias generales y la variabilidad en esas ventas.

Además, la estadística es fundamental para la toma de decisiones informadas. En un mundo donde la incertidumbre es la norma, la estadística nos proporciona herramientas para evaluar riesgos y oportunidades de manera objetiva. A través de la inferencia estadística, podemos tomar muestras representativas de poblaciones más grandes y generalizar nuestros hallazgos con cierto grado de confianza. Esto es crucial en campos como la medicina, donde los ensayos clínicos utilizan estadísticas para determinar la eficacia de nuevos tratamientos.

La estadística también juega un papel esencial en la

investigación científica y en la validación de hipótesis. Al aplicar métodos estadísticos a los datos recopilados en experimentos y estudios, los investigadores pueden determinar si los resultados observados son realmente significativos o simplemente el resultado del azar. Esto ayuda a garantizar la integridad y la calidad de la investigación en diversas disciplinas.

Además, la estadística es una herramienta esencial en el análisis y la interpretación de encuestas y estudios de opinión pública. A través de técnicas como la regresión y el análisis de varianza, podemos comprender cómo diferentes variables se relacionan entre sí y cómo influyen en los resultados. Esto es vital para comprender las preferencias de los consumidores, las tendencias sociales y políticas, y para tomar decisiones informadas basadas en las opiniones y actitudes de la población.

En general, las estadísticas se utilizan para analizar los datos de uno o más bloques de información, para procesarlos matemáticamente y obtener resultados con el fin de realizar diversas acciones en consecuencia. Se pueden destacar dos características:

- **Descripción de datos:** Esta función se relaciona con la representación gráfica de algunos resultados, haciéndolos más fáciles de entender.

- **Procesamiento de Datos:** Transforma el análisis teórico de los datos en un elemento práctico al proponer acciones que pueden ser implementadas en la realidad.

En el área de negocio, sirve para tener una visión integral de la situación actual de los procesos que se llevan a cabo, de manera que se puedan realizar cambios desde su administración. Además, permite descubrir las relaciones causa-efecto entre los diferentes componentes de un proceso productivo o financiero y detectar anomalías y variaciones en su ejecución.

1.1.2 Importancia de las estadísticas

La importancia de la estadística en el mundo contemporáneo es innegable, ya que esta disciplina no solo proporciona una comprensión profunda de los datos, sino que también impulsa la toma de decisiones informadas en una amplia gama de campos. Desde la ciencia hasta los negocios, desde la medicina hasta la política, la estadística desempeña un papel fundamental en la búsqueda de la verdad, la identificación de patrones y la predicción de resultados futuros.

En primer lugar, la estadística permite la exploración y el análisis sistemático de datos. Vivimos en una era de información abundante, y la estadística nos brinda las herramientas para ordenar, resumir y comprender esta avalancha de información. Mediante la creación de gráficos, tablas y medidas resumen, la estadística transforma datos crudos en información valiosa que puede ser interpretada y utilizada para extraer conocimiento significativo.

Además, la estadística desempeña un papel vital en la

investigación científica y la validación de hipótesis. Los científicos utilizan métodos estadísticos para determinar si los resultados de sus experimentos son significativos o simplemente el resultado del azar. Esto asegura la rigurosidad y la objetividad en el proceso científico, al tiempo que brinda confianza en la validez de los descubrimientos realizados.

• En el ámbito empresarial, la estadística es esencial para la toma de decisiones estratégicas. Las empresas recopilan grandes cantidades de datos sobre ventas, clientes, operaciones y más. Al analizar estos datos utilizando técnicas estadísticas, las empresas pueden identificar tendencias del mercado, segmentos de clientes y oportunidades de crecimiento. La estadística también es esencial en la gestión de riesgos y en la evaluación de la viabilidad de proyectos y productos nuevos.

• En la medicina, la estadística se utiliza para evaluar la eficacia de tratamientos y medicamentos. Los ensayos clínicos emplean métodos estadísticos para determinar si un nuevo tratamiento tiene un efecto real o si los resultados observados son simplemente el resultado del azar. Esto garantiza que los médicos y los pacientes puedan confiar en la eficacia y seguridad de los tratamientos recomendados.

• En la esfera política y social, la estadística es fundamental para comprender las opiniones y actitudes de la población. Encuestas y estudios utilizan técnicas

estadísticas para obtener resultados representativos y confiables. Esto permite a los líderes políticos y a los formuladores de políticas tomar decisiones informadas que reflejen las necesidades y deseos reales de la sociedad.

La estadística es el puente que conecta los datos con la comprensión y la acción. Su importancia radica en su capacidad para revelar patrones, proporcionar conocimientos y respaldar decisiones fundamentadas en la evidencia. En un mundo cada vez más impulsado por la información, la estadística es una herramienta esencial para navegar por la complejidad de los datos y transformarlos en conocimiento valioso que puede impulsar avances científicos, mejoras empresariales y decisiones políticas efectivas.

1.1.3 Objetivos estadísticos

El objetivo principal de la estadística es recopilar y procesar datos para que la información proporcionada por la investigación pueda aplicarse de manera práctica en soluciones, decisiones y otras acciones. Entre otros objetivos podemos mencionar:

• Proporcionar información para la toma de decisiones.

• Reducir posibles desperdicios o costos en diferentes procesos.

• Cuantificar un fenómeno para conocer su situación actual mediante la recopilación de datos diarios, semanales o mensuales.

• Comprobar el cumplimiento de objetivos.

• Cuantificar las características de un fenómeno para encontrar el promedio que impulsa el comportamiento de una población.

• Determinar la variabilidad de un fenómeno en el tiempo a través de la observación continua.

• Determinar las causas que producen un fenómeno.

• Comparar dos o más conjuntos de datos para determinar la existencia de una relación entre dos o más características de la misma clase.

• Hacer predicciones o pronósticos sobre el comportamiento de una población determinada.

• Extender las conclusiones del análisis de una muestra a toda una población.

• Facilitar la implementación de herramientas que determinen la validez y confiabilidad de los resultados.

• Facilitar la interpretación y comprensión de los estudios.

1.1.4 Características de la estadística

La estadística se caracteriza principalmente por el análisis matemático mediante el uso de expresiones aritméticas para representar numéricamente ciertas características pertenecientes a unos datos y facilitar su interpretación a través de diversos gráficos.

• Algunos de ellos explican grandes bloques de información.

• Se aplica a todas las áreas de la vida humana.

• Proporcionamos métodos de investigación para observar y recopilar toda la información necesaria para la investigación.

• Tiene un carácter especulativo ya que facilita la definición de previsiones a medio y largo plazo.

• Precisión de las conclusiones y resultados.

• Tiene un carácter tanto teórico como práctico, ya que permite realizar determinadas acciones sobre el objeto de estudio a partir del análisis.

• Es muy útil porque proporciona una forma de representar datos numéricos gráficamente.

1.2 Conceptos básicos

Cuando se utilizan métodos estadísticos en la investigación, normalmente se tienen en cuenta los siguientes factores en la recopilación y el procesamiento de datos:

• **Datos:** son los valores observados de la variable.

• **Unidades Muestrales:** Son los objetos de interés en un estudio. Las unidades de muestra pueden ser tornillos, personas o latas de frijoles como individuos, pero también pueden ser unidades formadas por muchos individuos,

como B. ciudades, escuelas o muchos tornillos.

• **Población de estudio:** el conjunto completo de unidades muestrales de interés para el estudio para responder a una pregunta de investigación. Por ejemplo, si estudiamos las orquídeas de un santuario, todas las orquídeas coincidirán con la población estudiada.

• **Muestreo de Población:** Cuando se quiere estudiar una población pero se encuentra con dificultades para acceder a todos sus miembros, se toma un subconjunto o subconjunto conocido como muestra. Por ejemplo, la ciudad de Ambato tiene más de 5 millones de habitantes. Si queremos estimar la estatura de los jóvenes entre 14 y 16 años en la ciudad de Ambato, tenemos que tomar una muestra de esta población, probablemente entre 100 y 500 personas.

• **Parámetro:** es el valor que describe una población.

• **Estadística:** es el valor determinado a partir de una muestra.

• **Error de estimación:** es la diferencia entre el estadístico muestral y el parámetro poblacional.

• **Margen de error:** Mide la diferencia máxima que se espera el 95 % del tiempo entre un resultado obtenido de una muestra y su valor poblacional real.

• **Muestreo:** Es la forma en que se seleccionan las unidades de muestreo para un estudio poblacional.

El muestreo puede ser un muestreo aleatorio simple, donde todos tienen las mismas posibilidades de ser seleccionados, o un muestreo sesgado, donde se seleccionan ciertos tipos de personas.

• **Muestreo Sistemático:** Comienza con una unidad elegida al azar, y de ahí, se toma una unidad por cada número específico especificado. Por ejemplo, desea realizar un estudio de control de calidad en una fábrica de salsa de tomate. Se selecciona una botella de salsa de tomate por cada 50 botellas de un lote de 5.000. Estas 100 botellas se someten a las pruebas necesarias para garantizar que el lote se encuentra en las condiciones adecuadas para la venta.

• **Muestra aleatoria estratificada:** La población se divide en grupos homogéneos llamados estratos, y de cada estrato se extrae una muestra aleatoria simple. Por ejemplo, en una ciudad se podría estratificar la población en menores de 10 años, entre 11 y 20 años, entre 21 y 30 años y mayores de 31 años.

• **Muestra representativa:** Una buena muestra debe ser representativa de la población. Esto significa que todas las características importantes de la población deben estar presentes en la muestra en la misma proporción que en la población.

• **Promedio:** es el valor característico o central de un conjunto de números. Dentro de estos valores, podemos hablar de la media (que se obtiene sumando todos los

valores del conjunto de datos y dividiendo la suma por la cantidad de datos de ese conjunto) y la mediana (el valor central del conjunto de datos ordenado).

• **Desviación estándar:** esta es una medida de dispersión basada en la media. Representa una distancia típica desde cualquier punto del conjunto de datos hasta su centro (medida por la media).

• **Intervalos de confianza (IC):** se refiere a un rango de valores posibles y un cierto porcentaje en el que se puede garantizar encontrar el valor verdadero de un parámetro poblacional. Por ejemplo, si el intervalo de confianza del 95% de la altura de una población es (165,175), eso significa que estará entre los valores 165 y 175 si repetimos las medidas el 95% de las veces.

• **Variables:** son características que pueden cambiar de una unidad de muestreo a otra. Hay variables numéricas (o cuantitativas) y variables categóricas (o cualitativas). Por ejemplo, la siguiente tabla muestra las variables de 10 perros llevados a un veterinario para consulta. Las variables género, raza y presencia de enfermedad son variables categóricas, mientras que la edad y el peso son variables numéricas.

1.3 Escalas de medición

Una escala de medida es un criterio de ordenamiento utilizado en estadística para organizar, clasificar y comparar conjuntos de datos. Son un sistema de clasificación en el que se puede ordenar la información

según una jerarquía predefinida.

Se puede definir como la forma en que algunos datos se relacionan y clasifican entre sí, de modo que, durante el análisis, se pueden ordenar de menor a mayor (o viceversa) y buscar similitudes entre variables.

En el corazón de la estadística y el análisis de datos yace un concepto esencial: las escalas de medición. Estas representan el sistema de valores y categorías utilizadas para cuantificar y categorizar las características de un objeto de estudio. Desde la escala más simple hasta la más compleja, las diferentes escalas proporcionan información vital sobre la naturaleza de los datos y las interpretaciones que podemos extraer de ellos. Aplicación de escalas de medida estadística

En diferentes análisis estadísticos, se tiene en cuenta una gran cantidad de datos que deben evaluarse con precisión. Pues la evaluación se basa en la comparación entre ellos, para determinar diferentes parámetros estadísticos, como frecuencia absoluta, tendencia y otros.

Las escalas de medida estadística permiten realizar estas comparaciones de forma eficaz, ya que ofrecen al investigador un sistema o criterio a partir del cual, según el tipo de datos, se pueden ordenar las distintas variables que intervienen en el análisis.

La mayoría de las aplicaciones de esta escala se basan en medir ciertas características, para determinar cuál es

mayor, cuál es menor, cuáles son iguales o diferentes, y operarlas matemáticamente.

1.3.1 Tipo de escala de medida

Las escalas de medida se clasifican en diferentes tipos, según el tipo de variables involucradas en la investigación y la forma en que se pueden jerarquizar los datos. También hay que tener en cuenta que cada grupo tiene ciertas operaciones.

Exploraremos cuatro tipos principales de escalas de medición: nominal, ordinal, de intervalo y de razón, y cómo influyen en la comprensión y el análisis de datos en diversas disciplinas.

- **Escala nominal**

En el nivel más básico de medición se encuentran las escalas nominales. Estas etiquetan o categorizan elementos, pero no establecen ningún orden o jerarquía entre ellos. Ejemplos incluyen el género, la afiliación política o la clasificación de especies. Aquí, las categorías son mutuamente excluyentes y exhaustivas, pero carecen de magnitud. No se pueden realizar operaciones matemáticas en estas escalas, ya que no tienen un punto de referencia absoluto. Son herramientas útiles para el análisis de la frecuencia y la proporción de categorías.

La escala nominal es el criterio a partir del cual se componen los datos cualitativos nominales, es decir, las cualidades o características que no tienen una jerarquía

determinada entre sí y, por tanto, no pueden regularse.

Se dice que las variables de este tipo son mutuamente excluyentes, una expresión que se ajusta a la lógica, indicando que dos o más estados no pueden existir simultáneamente. Es decir, en un determinado momento o contexto, dos eventos no pueden ocurrir simultáneamente.

Las únicas operaciones matemáticas que se pueden realizar para este tipo de escala son la igualdad (=) o, en su defecto, la diferencia (≠). Dado que los datos no se pueden ordenar, solo se pueden hacer comparaciones entre ellos utilizando las operaciones anteriores.

Ejemplo: sexo, código postal, estado civil, número telefónico, número al correr en un maratón, deporte favorito, carrera a estudiar, etc.

- **Escala ordinal**

Las escalas ordinales van un paso más allá al permitir la ordenación o jerarquización de los elementos, pero no establecen la magnitud de las diferencias entre ellos.

Ejemplos incluyen la clasificación de la satisfacción del cliente en términos de "insatisfecho", "neutral" o "satisfecho". Aunque podemos establecer un orden, no podemos decir con certeza cuánto mejor es "satisfecho" en comparación con "neutral". Las operaciones matemáticas limitadas, como medianas y percentiles, son aplicables a estas escalas.

La escala ordinal se utiliza para organizar datos cualitativos ordinales, es decir, aquellas características que tienen un orden específico. En general, el criterio de ordenación de este tipo de variables ha sido fijado artificialmente o se basa en otro criterio.

Al usar una escala ordinal, se pueden realizar las operaciones básicas de comparación, es decir, igualdad (=), desigualdad (≠), mayor que (>) y menor que (<).

Ejemplo: Pésimo – Malo – Regular – Bueno – Excelente

Primaria – Secundaria – Preparatoria – Licenciatura

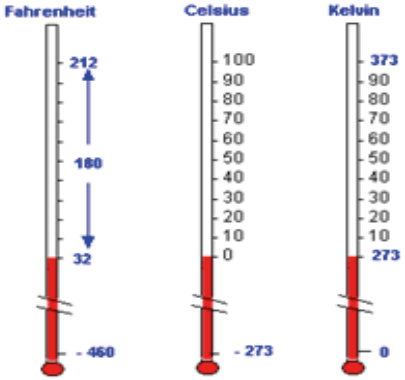
• **Escala de Intervalo**

La escala de intervalo es un tipo de escala de medida para organizar y trabajar con datos cuantitativos, es decir, datos numéricos. En este caso, la diferencia o distancia entre cada dato se conoce porque es constante.

Con escalas de este tipo se utiliza un punto cero arbitrario, es decir, un punto cero que no indica la ausencia de un valor, sino el punto más bajo que puede tomar una variable. Este punto cero se determina de mutuo acuerdo dentro del respectivo sistema de medición.

Con este tipo de escala, los datos se pueden sumar y restar, y comparar utilizando los operadores de comparación e igualdad.

Las escalas de intervalo permiten establecer relaciones de magnitud y diferencias iguales entre valores, pero no tienen un punto de inicio absoluto. La temperatura Celsius es un ejemplo clásico. Si bien podemos decir que 20°C es más caliente que 10°C y la diferencia es igual, no podemos afirmar que 20°C tiene el doble de calor que 10°C. Las operaciones matemáticas como sumar y restar son válidas, pero la multiplicación y la división no lo son debido a la carencia de un cero absoluto.



Ejemplo: Escalas de temperatura, la edad de la Tierra, la línea del tiempo de la humanidad.

- **Escala de razón**

Para datos cuantitativos se utiliza una escala de razón que no acepta valores menores a cero. Por lo tanto, se dice que este tipo de escala de medición implica el cero absoluto. Cero en este caso indica que no hay valor, a diferencia de la escala de intervalo.

Las dimensiones físicas como la velocidad, la distancia, la altura, el peso y la energía forman parte de este tipo de escala. Esto se debe a que cero significa que este estado físico no existe. De igual forma, no se aceptan valores negativos.

Este tipo de escala le permite realizar todas las operaciones aritméticas básicas (división, multiplicación, resta, suma) así como comparar datos usando operadores de comparación.

Las escalas de razón son las más completas, ya que permiten establecer relaciones de magnitud, diferencias iguales y un punto de partida absoluto. Ejemplos incluyen altura, peso y tiempo. En una escala de razón, podemos decir que un objeto es el doble de pesado que otro y que un valor de cero indica ausencia completa. Todas las operaciones matemáticas son aplicables en estas escalas, lo que las convierte en la forma más rica y versátil de medición.

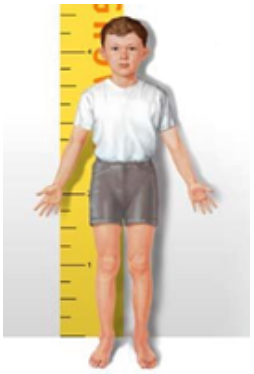
La elección de la escala de medición adecuada es esencial para garantizar la interpretación correcta de los datos y la aplicabilidad de las técnicas estadísticas. Cada tipo de escala impone limitaciones y posibilidades únicas en términos de análisis y generalización. Un estudio que no considere correctamente la escala de medición podría llevar a conclusiones erróneas o malinterpretaciones.

Como ejemplo, consideremos un estudio de opinión pública que evalúa la preferencia por diferentes sabores de helado utilizando una escala nominal. Si concluimos que el "sabor a vainilla" es más popular solo porque tiene la categoría más alta de frecuencia, estaríamos ignorando que las escalas nominales no tienen un orden real. Sin embargo, si aplicamos una escala de razón, podríamos determinar no solo qué sabor es más popular, sino también en qué grado es más preferido en comparación con otros sabores.

En el ámbito científico, la elección de la escala de

medición es crucial para garantizar la precisión de los resultados. Por ejemplo, en la investigación médica, la elección de la escala adecuada al medir variables como la intensidad del dolor puede afectar la interpretación de la efectividad de un tratamiento. Si se utiliza una escala ordinal para medir el dolor en lugar de una escala de razón, podríamos perder información valiosa sobre la magnitud de la mejora.

Ejemplo: peso, estatura, edad, distancia, dinero, etc.



- **Escalas comparativas y escalas no comparativas**

La investigación de mercado a menudo utiliza varios métodos y herramientas para la adquisición y el análisis de datos. Por este motivo, se han desarrollado diversas métricas para conocer qué opinan los clientes potenciales sobre los productos y servicios. Vale la pena mencionar que algunos de ellos también se utilizan en otros campos científicos.

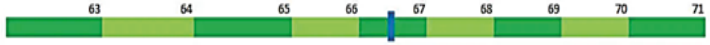
En este contexto, las escalas se dividen en escalas comparativas y escalas no comparativas.

Medida Comparativa: Métrica que te permite comparar las respuestas de los usuarios ya que existe una jerarquía que te permite conocer la satisfacción del público en general.

Escalas no comparativas: este tipo de escala no ordena los datos porque su propósito es recoger información sobre diferentes temas de estudio, no compararlos entre sí.

- **Escala gráfica de medición de datos**

Una escala gráfica de medición de datos es una representación visual de una escala ordinal, de intervalo o de razón, donde diferentes variables se ubican en una recta numérica. De esta forma, los datos se pueden ubicar en cualquier punto de la línea, por lo que se considera un valor continuo, es decir, puede tomar cualquier valor.



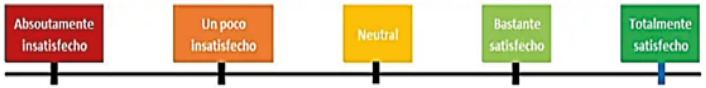
Ejemplo: La persona, como se observa, tiene un peso actual de 66,5 kilos. Esta forma de representar este tipo de datos permite interpretar, con mayor precisión, una situación.

- **Escala Likert**

La escala de Likert es un tipo de escala, desarrollada por el psicólogo Rensis Likert, que se utiliza para medir el nivel de satisfacción de una persona con un objeto o situación. En este caso, el encuestado puede elegir una

de varias opciones (normalmente cinco) que se ordenan según su nivel de optimismo.

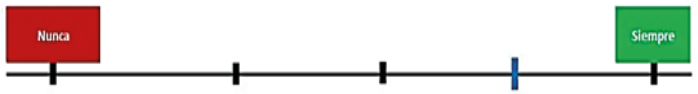
Los extremos, en este caso, están relacionados con la insatisfacción total por parte del consumidor y la satisfacción absoluta. Entre estos límites existen estados más o menos apropiados (p. ej., "ligeramente satisfecho", "moderadamente insatisfecho", "satisfecho").



- **Escala diferencial semántica (Max Diff)**

La escala de diferencial semántico es un sistema de calificación que permite conocer el comportamiento o actitud de una persona en respuesta a un objeto. En este aspecto es similar a la escala de Likert, sin embargo, las variables o grados no se reducen al grado de satisfacción de los individuos, sino que se extienden al análisis de diversas opiniones.

En esta escala solo se indican sus valores extremos, del más negativo al más positivo. Sin embargo, muchas veces se permite al usuario seleccionar opciones que son intermedias pero que no tienen una etiqueta específica, como es el caso de la escala Likert.



- **Escala matricial para análisis de datos paralelos**

La escala de análisis de datos paralelos Matrix se utiliza para comparar la importancia percibida y la satisfacción de los usuarios con un producto. Por lo tanto, los investigadores pueden tomar acciones específicas basadas en los resultados de la encuesta.

Por ejemplo, si un elemento se considera importante pero la satisfacción del usuario es baja, se pueden implementar estrategias para mejorar esa percepción.

	Importancia					Satisfacción				
	Sin importancia				Con importancia	Insatisfecho				Satisfecho
	1	2	3	4	5	1	2	3	4	5
Sitio web					X		X			
Transporte			X					X		
Productos					X					X

Las escalas de medida son una base importante para la correcta aplicación del método estadístico en cualquier ámbito profesional, científico o académico. Tanto en la estadística descriptiva como en la de inferencia, son una herramienta fundamental para organizar los datos, compararlos entre sí y extraer conclusiones adecuadas para conocer con exactitud la naturaleza del objeto de estudio.

Actualmente, existen diversas herramientas informáticas para desarrollar este tipo de trabajo analítico, entre las que destaca Microsoft Excel. Gracias a sus funciones estadísticas y diversas posibilidades para organizar los datos, Excel se ha convertido en una

herramienta eficaz para los ejercicios estadísticos.

1.4 Representación gráfica

Es una forma eficiente de descubrir el comportamiento de un conjunto de datos porque permite una descripción rápida y fácil de entender. Su importancia es tal que cualquier análisis estadístico debe ir acompañado de esta forma.

1.4.1 Puntos

Es un método para resumir datos cuantitativos en el que cada observación está representada por un punto en una recta numérica. Si tiene un resumen de muchos datos, cada punto puede representar un número fijo de personas. Este gráfico muestra la ubicación general de las observaciones, su dispersión y la presencia de observaciones inusuales, atípicas o extremas. Se recomienda usarlo cuando se muestran un máximo de 20 observaciones individuales, de lo contrario es difícil distinguirlas.

Puede combinar sus gráficos de dos o más conjuntos de datos en el mismo gráfico con un método simple de interpretación. Por ejemplo: triángulos, círculos, cuadrados, rectángulos o cualquier otra forma en lugar de puntos.

La creación de este tipo de gráfico determinará el rango de observaciones, cómo se trazarán y una escala adecuada que permitirá que sus datos se representen

bien. Para datos nominales u ordinales, un gráfico de dispersión es similar a un gráfico de barras, pero se reemplaza con una serie de puntos. Para datos continuos, este gráfico se asemeja a un histograma, con rectángulos reemplazados por puntos.

1.4.2 Tallos y hojas

El gráfico de dispersión tiene algunas desventajas, p. B. rastrear puntos hasta troncales y puede ser confuso con grandes cantidades de datos. Es conveniente utilizar otras herramientas gráficas. El diagrama de tallo y hojas es una técnica semigráfica que se utiliza para ilustrar las características clave de los datos, como la ubicación, la dispersión y la simetría. Tiene la ventaja de representar valores de datos y por su forma puede usarse para conjuntos de datos de hasta 100 elementos.

Ejemplo:

08	19	17	01	07	09	05	16
13	15	04	02	00	04	01	12

Los datos se clasifican consideran las decenas tal que se consideran dos grupos. Uno comienza con 0 y el otro en 1 formando el tallo verticalmente:

0	
1	

2) Para cada elemento se anota el segundo dígito a la derecha de barra vertical, que construyen las hojas:

0	8	1	7	9	5	4	2	0	4	1
1	9	7	6	3	5	2				

Se ordena en forma ascendente:

0	0	1	1	2	4	4	5	7	8	9
1	2	3	5	6	7	9				

Se crean dos categorías en forma ascendente en cada decena, los dígitos de unidades del 0-4 forman el primero y 5-9 forman el segundo grupo:

0	0	1	1	2	4	4
	5	7	8	9		
1	2	3				
	5	6	7	9		

En casos en que la base de datos consta de más de dos cifras se escoge los rangos par agrupaciones que se harán. Después, mediante una coma se separan, llenadas las hojas:

33	55	79	106	188	47	118	248
47	58	82	113	208	60	88	

Con base en estos datos, se puede construir dos diagramas de tallo y hojas:

1) Primer diagrama:

Categoría	Dígito									
0-100	0	33	47	47	55	58	60	79	82	88
100-200	1	06	13	18	88					
200-300	2	08	48							

O, también:

Categoría	Dígito						
0-50	0	33	47	47			
50-100	0	55	58	60	79	82	88
100-150	1	06	13	18			
150-200	1	88					
200-250	2	08	48				
250-300	2						

Sin embargo, diagramas múltiples se pueden usar para comparar dos conjuntos de datos. Por consiguiente, se coloca un tallo común y hojas de un conjunto se sitúan a la izquierda y hojas del segundo conjunto a la derecha del tallo, respectivamente:

				Dígito						
			44	1	33	47	47			
		57	79	1	55	58	60	79	82	88
01	23	34	42	2	06	13	18			
		06	78	2	88					
			33	3	08	48				
			5	3						
				4						

Los datos izquierdos están más agrupados en respecto a valores bajos, con rango mayor y fuerte asimetría, mientras que el conjunto derecho es asimétrico y con menor dispersión. Finalmente, estos diagramas se emplean para representar datos con decimales:

0.80	0.46	1.23	1.15	2.23	1.89	0.95	1.02	2.06	0.61	0.52	1.94

Su diagrama ascendente es:

Categoría	Dígito					
0–1.0	0.	46	52	61	80	95
1.0–2.0	1.	02	15	23	89	94
2.0–3.0	2.	06	23			

1.4.3 Sectores circulares o gráfica de pastel

Una alternativa para mostrar las frecuencias relativas de un conjunto de categorías es utilizar gráficos circulares. En este caso, a cada categoría se le asigna un sector que representa su frecuencia. Una limitación de este tipo de representación es el número de categorías, ya que cuando son muchas, la lectura del gráfico no es fácil.

Figura 1. Representación gráfica de sectores de pastel

1.4.4 Histograma

Es una serie de rectángulos, cada uno de los cuales representa un intervalo de agrupación. Sus bases corresponden al intervalo de clase utilizado en la distribución de frecuencias y las alturas son proporcionales a la frecuencia absoluta n_i o relativa f_i de la clase. Es adecuado para datos continuos medidos a la misma escala y se utiliza cuando la creación de un gráfico de tallo-hoja requiere mucha mano de obra. Puede ayudar a descubrir valores atípicos y brechas entre los datos.

Otra forma alternativa de presentar los resultados de la figura 2 el histograma clásico. La siguiente figura muestra el histograma de las abundancias relativas y el polígono correspondiente a los puntajes de los estudiantes. Lo más destacable que se puede observar es la marcada asimetría de la distribución; En comparación con la representación de diagrama de caja, es más difícil identificar los percentiles.

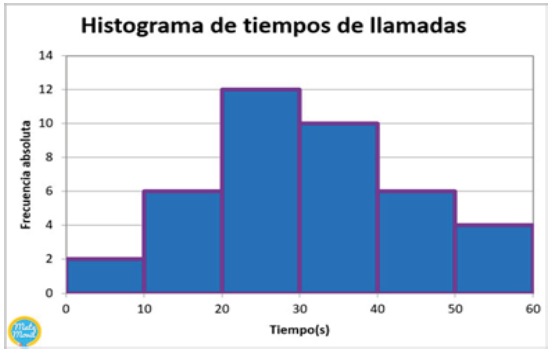


Figura 2. Representación gráfica del histograma

1.4.5 Polígono de frecuencia

Este es un gráfico creado al conectar segmentos de línea a puntos que tienen la marca de clase proporcional en la abscisa y la frecuencia correspondiente en la ordenada. Cierra en ambos extremos en marcas adyacentes con frecuencia cero.

Cuando estás examinando la relación entre dos variables (X e Y, por ejemplo), es muy útil crear un diagrama de dispersión. Esta es una gráfica donde cada observación está representada en el plano XY por un punto cuyas coordenadas están dadas por los valores registrados en ambas variables. En algunos casos, un diagrama de dispersión se puede modificar insertando segmentos de línea que conectan los puntos del plano según un orden dado por el eje de abscisas. A modo de ejemplo, supongamos que se evalúa el número de publicaciones pér cápita obtenidos por un grupo de docente por año .

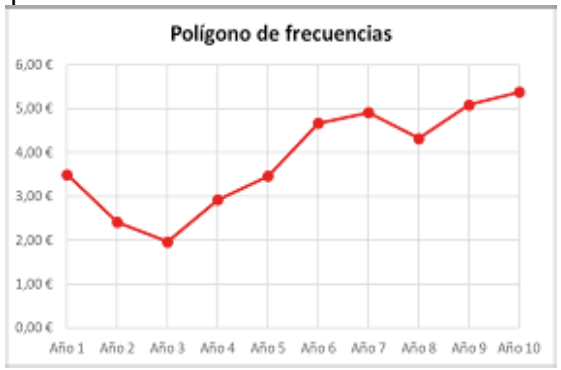


Figura 3. Representación gráfica del polígono de frecuencias

1.4.6 Ojiva

En estadística, una ojiva es un gráfico que muestra la curva de una función de distribución acumulativa, dibujada a mano o en un software de computadora. Los puntos trazados son el límite de clase superior y la frecuencia acumulada correspondiente. La ojiva para la distribución normal se asemeja a un lado de un arabesco u arco ojival. El término también se puede utilizar para la función de distribución acumulativa empírica. Este es un tipo de gráfico de frecuencia y también se conoce como polígono de frecuencia acumulada. Se utiliza para indicar el número (o proporción) de observaciones que son menores o iguales a un cierto valor.

Una ojiva se construye sobre un sistema de ejes perpendiculares. En el eje horizontal ingresamos los límites de los intervalos de clase previamente determinados. En función de cada uno de estos umbrales, determinamos en la ordenada la altura que corresponde a la frecuencia acumulada correspondiente a ese valor. Si conectamos los puntos sucesivos así determinados por segmentos de recta, obtenemos una recta llamada ojiva. Por lo general, representamos la frecuencia acumulada en el eje vertical izquierdo y el porcentaje de frecuencia acumulada en el eje vertical derecho.

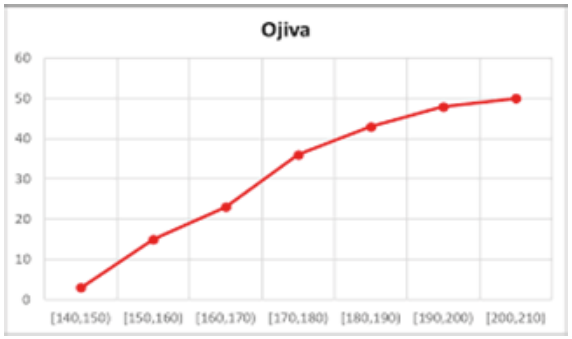


Figura 4. Representación gráfica de la ojiva

1.5 Tipos de Estadística

Las estadísticas se dividen en varias categorías según el propósito para el que se aplican y los métodos utilizados. Sin embargo, cabe señalar que cada uno de sus tipos es parte fundamental de cualquier campo de investigación y son interdependientes. Este tipo de estadísticas son:

1.5.1 Estadísticas descriptivas

Las estadísticas descriptivas se refieren a ramas que tienen información para organizarlas a través de gráficos y tablas, y describirlas numéricamente a través de varias fórmulas. Todo esto se hace para facilitar su interpretación y comprensión, ya sea con fines educativos o informativos o en entornos donde diferentes investigadores necesitan acceder a la información rápidamente.

1.5.2 Estadísticas inferencial

La estadística inferencial, por su parte, es un enfoque en el que se determinan las predicciones probables sobre

el comportamiento futuro de un grupo de datos, de ahí el nombre, ya que se infieren determinados resultados. Este tipo de estadística está relacionada con la probabilidad. A su vez, se divide en dos subtipos:

- **Estadísticas paramétricas:** En este tipo de estadísticas, un conjunto de datos presenta una determinada distribución de probabilidad, por lo que se pueden aplicar ciertos parámetros a su análisis.

- **Estadísticas no paramétricas:** en este caso, el grupo de datos no se puede analizar con ciertos parámetros, porque tampoco se puede identificar la distribución.

1.5.3 Estadísticas aplicadas

En esta área de la estadística, la información que se obtiene de un estudio muestral se extiende a toda la población de la que se extrajeron, de modo que se explican todos los elementos.

1.5.4 Estadística matemática

Se refiere a un enfoque estadístico en el que se aplican diferentes fórmulas matemáticas para obtener resultados precisos y concretos del grupo de datos que se analiza.

1.6 Campos de aplicación y áreas de estadística

La estadística como método científico se puede aplicar en muchas áreas donde los datos cuantitativos y cualitativos deben analizarse juntos. Por un lado, su

carácter teórico permite su uso en el ámbito académico de la educación.

Un área donde se destaca esta adopción es en el área de negocios. Aplicación, tanto en términos de la capacidad de proporcionar grupos de investigación para la validación inmediata de los logros de la empresa, como en el tiempo para crear soluciones oportunas basadas en las predicciones obtenidas de las empresas. En este caso, esta es una herramienta útil si:

• **Procesos de producción:** Deben introducirse técnicas estadísticas para evaluar la calidad y la eficacia de los procesos de producción para excluir, cambiar o mejorar los procesos de producción que tienen un impacto directo en la productividad de una empresa en comparación con la aplicación de indicadores clave de desempeño.

• **Finanzas:** El análisis estadístico le permite rastrear y registrar varias áreas financieras de su empresa, como presupuestos, gastos e inversiones, para que pueda observar el comportamiento de su empresa y tomar decisiones correctas y precisas.

• **Contabilidad:** Al optimizar los gastos y determinar la posición financiera de la empresa, las estadísticas pueden determinar mejoras y soluciones que se pueden implementar desde una perspectiva contable.

• **Capital Humano:** El sector de recursos humanos necesita realizar investigaciones permanentes sobre

el clima laboral y las tendencias que muestren ciertos aspectos que afectan la calidad del trabajo.

• **Marketing:** La investigación es una herramienta estadística aplicada en esta área para evaluar la demanda y el potencial de aumento o disminución de la demanda para que se puedan proponer nuevos productos.

1.7 Etapas de los métodos estadísticos.

Como todo proceso de investigación, todo método estadístico implica pasos que se deben seguir para su adecuada implementación. Estas etapas son:

• **Recopilación de datos:** En esta etapa se identifica la población a analizar y se decide qué parte o muestra de esa población se estudiará. Se utilizan varios métodos de recopilación de datos, como encuestas.

• **Análisis matemático de datos:** una vez que se recopilan los datos, se aplican varias fórmulas matemáticas para representar, describir y resumir la información numéricamente.

• **Descripción y presentación de la información:** Los resultados obtenidos de los distintos cálculos realizados se organizan en gráficos y tablas para su fácil comprensión.

Análisis comparativo y conclusiones: Finalmente, el grupo de investigación se reúne para comparar y analizar los resultados, extraer conclusiones y evaluar

posibles medidas o acciones que se puedan aplicar a la situación o fenómeno investigado.

CAPÍTULO II

TEORÍA DE PROBABILIDADES

La teoría de probabilidades es una disciplina fundamental que nos permite abordar la incertidumbre inherente en muchos aspectos de la vida y de la ciencia. Desde la predicción del clima hasta la toma de decisiones en los negocios, la probabilidad desempeña un papel crucial al ayudarnos a cuantificar y comprender el riesgo y la variabilidad en diferentes situaciones. Esta rama de las matemáticas ha sido moldeada y refinada a lo largo de los años, brindando una base sólida para abordar fenómenos aleatorios y sus patrones subyacentes.

"La probabilidad es la guía de la vida." Esta afirmación del matemático Joseph Juran resalta la importancia de la probabilidad en nuestra toma de decisiones diarias y en la comprensión de eventos inciertos. La teoría de probabilidades se basa en la idea fundamental de que, en muchos casos, no podemos prever con certeza el resultado de un evento futuro, pero podemos asignar probabilidades a los diferentes resultados posibles. Esta asignación de probabilidades nos brinda una forma estructurada de enfrentar la incertidumbre, permitiéndonos tomar decisiones informadas.

La teoría de probabilidades tiene sus raíces en la correspondencia entre los matemáticos Blaise Pascal y Pierre de Fermat en el siglo XVII, que buscaban resolver problemas relacionados con juegos de azar. Esta correspondencia sentó las bases para el cálculo de probabilidades y la noción de que las posibilidades pueden ser evaluadas de manera cuantitativa. Además,

el matemático Abraham de Moivre y su teorema del límite central jugaron un papel crucial en la consolidación de la distribución normal, que se convirtió en una piedra angular de la estadística moderna.

La teoría de probabilidades no solo es esencial en matemáticas y estadísticas, sino que también encuentra aplicaciones en una amplia gama de disciplinas. En la economía, ayuda a modelar los mercados financieros y las inversiones. En la medicina, contribuye a la evaluación de riesgos y a la toma de decisiones clínicas. En la inteligencia artificial, forma la base de algoritmos de aprendizaje automático y procesamiento del lenguaje natural. Cada uno de estos campos se beneficia de la capacidad de la teoría de probabilidades para analizar y cuantificar la incertidumbre inherente en los datos y los procesos.

2.1 Definición técnica

La teoría de la probabilidad es una herramienta matemática que establece un conjunto de reglas o principios útiles para calcular la ocurrencia o no ocurrencia de fenómenos aleatorios y procesos estocásticos

En otras palabras, la teoría de la probabilidad consiste en todo el conocimiento relacionado con el concepto de probabilidad. Básicamente, es un concepto matemático. Asimismo, la probabilidad, como rama de las matemáticas, es una herramienta de la estadística.

En cualquier caso, sin desviarnos del concepto de teoría de la probabilidad, diremos que consiste en un conjunto de trucos que nos permiten asignar un número a la posibilidad de un evento.

Entonces, en el caso de las monedas, sabemos que cuando se lanzan al aire, el resultado puede ser cara o cruz. Suponiendo que la moneda y el tablero son perfectos y las condiciones para el lanzamiento no cambian, la probabilidad es del 50% de cara y del 50% de cruz.

Fue en este punto que nació el concepto de probabilidad. La probabilidad es un número entre 0 y 1, generalmente expresado como un % entre 0 y 100, que indica el número promedio de veces que ocurrirá un evento cada 100 veces.

Teniendo esto en cuenta, llegamos a la conclusión de que la teoría de la probabilidad se encarga de estudiar el número entre 0 y 1 que tenemos que asignar a un determinado evento. Es decir, se encarga de estudiar la probabilidad de que ocurra un evento.

2.2 Historia de la teoría de la probabilidad

La teoría de la probabilidad, enraizada en el intento humano de comprender y cuantificar la incertidumbre, ha evolucionado a lo largo de los siglos, moldeando tanto el pensamiento matemático como la toma de decisiones en diversos campos. Desde sus inicios en los juegos de azar hasta su aplicación en áreas como la estadística, la economía y la ciencia, la historia de la teoría de la

probabilidad es un testimonio del ingenio humano en la búsqueda de la comprensión de lo incierto.

La teoría de la probabilidad encuentra sus primeras raíces en la correspondencia entre Blaise Pascal y Pierre de Fermat en el siglo XVII. Aunque su objetivo inicial era resolver problemas relacionados con juegos de azar, sus discusiones sentaron las bases para la cuantificación de posibilidades y la asignación de probabilidades a resultados inciertos (Pascal & Fermat, 1654).

Abraham de Moivre desempeñó un papel fundamental en la consolidación de la teoría de la probabilidad en el siglo XVIII. Su trabajo en el teorema del límite central sentó las bases para la comprensión de cómo se acumulan los errores aleatorios en torno a un valor central. Además, su contribución a la distribución normal, también conocida como la curva de Gauss, estableció un modelo crucial para describir la variabilidad en una amplia gama de fenómenos (de Moivre, 1733).

En el siglo XX, la teoría de la probabilidad experimentó un renacimiento con la aparición de la teoría de la medida y la probabilidad moderna. Kolmogorov, en su influyente obra "Foundations of the Theory of Probability" (1933), estableció un marco axiomático riguroso que unificó la teoría de la probabilidad y la convirtió en una rama respetada de las matemáticas.

Hoy en día, la teoría de la probabilidad es esencial en campos que van desde la estadística y la econometría

hasta la inteligencia artificial y la toma de decisiones médicas. Ha dejado su huella en la comprensión de riesgos y en la construcción de modelos precisos para describir fenómenos aleatorios en diversas disciplinas (Ross, 2010).

2.3 Tipos de probabilidades

La teoría de la probabilidad es una disciplina esencial en matemáticas y estadísticas que aborda la cuantificación de la incertidumbre en diversos contextos. A lo largo de su desarrollo, han surgido varios enfoques y tipos de probabilidades que reflejan diferentes perspectivas sobre cómo entender y modelar eventos inciertos. En este documento, exploraremos los tres tipos principales de probabilidades: clásica, frecuentista y subjetiva.

2.3.1 Probabilidad Clásica

La probabilidad clásica, también conocida como probabilidad a priori, se basa en la noción de equiprobabilidad, donde todos los resultados posibles tienen la misma probabilidad de ocurrir. Este enfoque se aplica comúnmente a experimentos teóricos en los que se pueden identificar todos los resultados posibles y se asume que son igualmente probables.

Por ejemplo, en el lanzamiento de un dado justo de seis caras, cada resultado (1 al 6) se considera equiprobable. En este contexto, la probabilidad de obtener un número específico, como 3, sería 1/6.

La probabilidad clásica, también conocida como probabilidad a priori, es uno de los conceptos fundamentales en la teoría de la probabilidad. Se basa en la idea de que todos los resultados posibles de un experimento son igualmente probables cuando no hay razones para considerar lo contrario. En este documento, exploraremos en profundidad la probabilidad clásica, sus aplicaciones y ejemplos ilustrativos.

Conceptos Básicos de la Probabilidad Clásica

La probabilidad clásica se apoya en la suposición de que todos los resultados posibles de un evento tienen la misma probabilidad de ocurrir. En otras palabras, si un experimento tiene \(n\) resultados posibles igualmente probables, la probabilidad de que un resultado en particular ocurra es \(1/n\). Esto se debe a que cada resultado tiene la misma "oportunidad" de ser el resultado final.

- **Ejemplo 1: Lanzamiento de Moneda**

Un ejemplo clásico de probabilidad clásica es el lanzamiento de una moneda justa. Si la moneda está equilibrada y no hay razones para creer que caerá de manera diferente, los resultados posibles son "cara" y "cruz". Dado que ambos resultados son igualmente probables, la probabilidad de obtener "cara" es \(1/2\) y la probabilidad de obtener "cruz" también es \(1/2\).

- **Ejemplo 2: Lanzamiento de Dado**

Similarmente, consideremos el lanzamiento de un dado justo de seis caras. Cada cara tiene la misma probabilidad de aparecer debido a la simetría del dado. Por lo tanto, la probabilidad de obtener cualquier número específico (1 al 6) es \(1/6\).

Aplicaciones de la Probabilidad Clásica

La probabilidad clásica se aplica en una variedad de situaciones en las que los resultados son equiprobables. Aunque puede parecer simple, este enfoque tiene importancia práctica en áreas como la teoría de juegos, la educación y la toma de decisiones.

- **Ejemplo 3: Extracción de Bolas de una Urna**

Supongamos que tenemos una urna con bolas de colores: 3 bolas rojas, 2 bolas verdes y 5 bolas azules. Si seleccionamos una bola al azar sin mirar, la probabilidad de obtener una bola roja es \(3/10\), ya que hay un total de 10 bolas y 3 son rojas. De manera similar, la probabilidad de obtener una bola verde es \(2/10\) y la probabilidad de obtener una bola azul es \(5/10\).

- **Ejemplo 4: Distribución de Cartas**

En un mazo estándar de 52 cartas, la probabilidad de obtener cualquier carta específica es \(1/52\), ya que todas las cartas son consideradas igualmente probables cuando se baraja el mazo de manera adecuada.

La probabilidad clásica es un enfoque fundamental en la teoría de la probabilidad que se basa en la igualdad de probabilidades para todos los resultados posibles de un experimento. Aunque puede parecer simplista en comparación con enfoques más avanzados, la probabilidad clásica tiene aplicaciones importantes en diversas áreas y proporciona una base sólida para entender la incertidumbre en situaciones donde no tenemos información adicional.

2.3.2 Probabilidad Frecuentista

La probabilidad frecuentista se basa en la idea de que la probabilidad de un evento se puede calcular observando la frecuencia con la que ocurre en un gran número de repeticiones de un experimento. Este enfoque está arraigado en la estadística y se utiliza para modelar situaciones en las que no se conocen las probabilidades a priori.

Según este enfoque, si lanzamos un dado muchas veces y contamos la frecuencia con la que aparece un número determinado, la probabilidad de ese número será aproximadamente la relación entre las veces que apareció y el número total de lanzamientos.

La probabilidad frecuentista es uno de los enfoques fundamentales para comprender la probabilidad en el contexto de la teoría de la probabilidad. Se basa en la idea de que la probabilidad de un evento se calcula observando la frecuencia relativa con la que ese evento

ocurre en un número significativo de repeticiones de un experimento. En este documento, exploraremos en qué consiste la probabilidad frecuentista y proporcionaremos ejemplos resueltos para una comprensión más clara.

Definición de Probabilidad Frecuentista

Según Ross, S. M. (2019), la probabilidad frecuentista se define como "la proporción límite de veces que ocurre un evento en un número infinito de repeticiones del mismo experimento" (p. 21). En esencia, se basa en la noción de que la probabilidad es una medida de la tendencia observable de un evento a medida que se repite un experimento en un número significativamente grande de ocasiones.

- **Ejemplo 1**

Lanzamiento de un Dado

Supongamos que tenemos un dado de seis caras y queremos calcular la probabilidad de obtener un número impar. En un enfoque frecuentista, lanzamos el dado muchas veces y contamos cuántas veces obtenemos un número impar.

Lanzamos el dado 100 veces.

Obtenemos un número impar (1, 3 o 5) en 51 de esas ocasiones.

La probabilidad frecuentista de obtener un número impar es entonces 51/100 = 0.51.

- **Ejemplo 2**

Lanzamiento de Monedas

Consideremos el lanzamiento de una moneda justa. Queremos calcular la probabilidad de obtener cara. Lanzamos la moneda 500 veces y registramos cuántas veces cae cara.

Obtenemos cara en 250 ocasiones.

La probabilidad frecuentista de obtener cara es 250/500 = 0.5.

- **Ejemplo 3**

Enfermedades y Diagnóstico

Imaginemos que estamos estudiando la probabilidad de que un paciente tenga una cierta enfermedad dado un conjunto de síntomas. Si realizamos pruebas en un gran número de pacientes y contamos cuántos de ellos tienen la enfermedad y cuántos no, podemos estimar la probabilidad frecuentista de que un paciente con esos síntomas tenga la enfermedad.

La probabilidad frecuentista se basa en observar la frecuencia relativa de un evento a medida que se repite un experimento en un número significativo de ocasiones. A través de ejemplos resueltos, se ilustra cómo este enfoque se aplica en situaciones como lanzamiento de dados, monedas y análisis médicos.

2.3.3 Probabilidad Subjetiva

La probabilidad subjetiva, también conocida como probabilidad bayesiana, se basa en la creencia personal o el juicio subjetivo de un individuo sobre la probabilidad de que ocurra un evento. En este enfoque, la probabilidad se asigna en función de la información y las creencias disponibles en ese momento.

Este tipo de probabilidad es especialmente útil cuando se trata de situaciones en las que no se dispone de datos históricos o cuando la incertidumbre es alta. Se aplica en campos como la toma de decisiones, la economía y la medicina, donde los expertos utilizan su conocimiento y experiencia para asignar probabilidades a eventos futuros.

La probabilidad subjetiva es un enfoque único en la teoría de la probabilidad, ya que se basa en las creencias y juicios personales de un individuo. A diferencia de los enfoques frecuentista y clásico, que se basan en datos observables y reglas precisas, la probabilidad subjetiva considera la incertidumbre desde una perspectiva individual y se basa en la evaluación subjetiva de la probabilidad de un evento.

Definición de Probabilidad Subjetiva

La probabilidad subjetiva se basa en las creencias personales de un individuo sobre la probabilidad de que ocurra un evento particular. De acuerdo con de Finetti,

B. (1974), esta probabilidad "es una medida de la fuerza con la que una persona mantiene una creencia sobre la ocurrencia de un evento" (p. 293). En otras palabras, es la evaluación personal y subjetiva de la certeza o incertidumbre acerca de un resultado.

- **Ejemplo 1**

Elección de un Candidato

Imaginemos que un votante debe decidir entre dos candidatos en una elección. La probabilidad subjetiva de que el Candidato A gane podría ser del 60%, mientras que la probabilidad subjetiva de que el Candidato B gane podría ser del 40%. Estas evaluaciones no se basan en datos históricos ni en reglas matemáticas, sino en las creencias personales del votante.

- **Ejemplo 2**

Clima en un Viaje

Supongamos que una persona está planeando un viaje a la playa. Antes de partir, evalúa subjetivamente la probabilidad de que llueva durante el viaje. Si considera que hay un 30% de probabilidad de lluvia, está expresando su juicio personal basado en factores como el pronóstico del clima, las estaciones del año y su propia intuición.

- **Ejemplo 3**

Inversión en el Mercado de Valores

Un inversionista evalúa subjetivamente la probabilidad de que el precio de una acción aumente en un cierto período de tiempo. Esta evaluación se basa en su experiencia, investigación y conocimientos del mercado, y no en datos objetivos o fórmulas matemáticas.

2.4 Operaciones con probabilidades

Las operaciones con probabilidades son esenciales para el análisis y la toma de decisiones en diversos campos, desde la estadística hasta la teoría de juegos. Estas operaciones permiten combinar y manipular diferentes probabilidades para obtener información valiosa sobre eventos futuros. En este documento, exploraremos las operaciones comunes que podemos realizar con las probabilidades, proporcionaremos ejemplos resueltos.

2.4.1 Unión de Eventos (Unión)

La unión de dos eventos, A y B, se representa como "$A \cup B$" y significa que al menos uno de los eventos debe ocurrir. La probabilidad de la unión se calcula como la suma de las probabilidades individuales menos la probabilidad de su intersección.

Ejemplo:

Supongamos que en un torneo de baloncesto, la probabilidad de que el Equipo A gane es 0.6 y la probabilidad de que el Equipo B gane es 0.4. La probabilidad de que ambos equipos ganen sus partidos

respectivos es 0.2.

¿Cuál es la probabilidad de que al menos uno de los equipos gane?

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

- $P(A \cup B) = 0.6 + 0.4 - 0.2 = 0.8$

2.4.2 Intersección de Eventos (Intersección)

La intersección de dos eventos, A y B, se representa como "$A \cap B$" y significa que ambos eventos deben ocurrir al mismo tiempo. La probabilidad de la intersección se calcula multiplicando las probabilidades individuales.

Ejemplo:

Siguiendo con el ejemplo anterior, ¿cuál es la probabilidad de que ambos equipos ganen sus partidos respectivos?

- $P(A \cap B) = P(A) * P(B)$

- $P(A \cap B) = 0.6 * 0.4 = 0.24$

2.4.3 Complemento de un Evento

El complemento de un evento A, denotado como "A"", representa la no ocurrencia de A. La probabilidad del complemento es 1 menos la probabilidad del evento original.

Ejemplo:

Si la probabilidad de que llueva en un día es 0.3, ¿cuál

es la probabilidad de que no llueva?

- P(No llueva) = 1 - P(Llueva)

- P(No llueva) = 1 - 0.3 = 0.7

2.4.4 Eventos Mutuamente Excluyentes

Los eventos mutuamente excluyentes son aquellos que no pueden ocurrir simultáneamente. La probabilidad de la unión de eventos mutuamente excluyentes se calcula sumando las probabilidades individuales.

Ejemplo:

En un lanzamiento de un dado de seis caras, ¿cuál es la probabilidad de obtener un número par o un número impar?

- P(Número par o impar) = P(Número par) + P(Número impar)

- P(Número par o impar) = 3/6 + 3/6 = 1

2.4.5 Regla de Inclusión-Exclusión

Esta regla se utiliza para calcular la probabilidad de la unión de más de dos eventos. Se basa en restar las probabilidades de las intersecciones de los eventos, sumar las probabilidades de las intersecciones de tres eventos y así sucesivamente.

Ejemplo:

En un conjunto de estudiantes, 40 estudian

Matemáticas, 30 estudian Física y 20 estudian Química. Si 10 estudian tanto Matemáticas como Física, 8 estudian tanto Física como Química, 5 estudian tanto Matemáticas como Química, y 2 estudian las tres materias.

¿Cuál es la probabilidad de que un estudiante estudie al menos una de estas materias?

- P(Matemáticas ∪ Física ∪ Química) = P(Matemáticas) + P(Física) + P(Química) - P(Matemáticas ∩ Física) - P(Física ∩ Química) - P(Matemáticas ∩ Química) + P(Matemáticas ∩ Física ∩ Química)

- P(Matemáticas ∪ Física ∪ Química) = 40/100 + 30/100 + 20/100 - 10/100 - 8/100 - 5/100 + 2/100 = 0.65

Las operaciones con probabilidades son fundamentales para comprender la incertidumbre y tomar decisiones informadas en diversos campos. Mediante ejemplos resueltos, hemos ilustrado cómo realizar cálculos de unión, intersección, complemento y más. Las citas bibliográficas respaldan la base teórica de estas operaciones y su aplicación en situaciones reales. La manipulación de probabilidades brinda una herramienta poderosa para evaluar escenarios y modelar eventos futuros en condiciones de incertidumbre.

2.5 Ejercicios de probabilidades

La teoría de probabilidades, aunque a menudo considerada una rama abstracta de las matemáticas, tiene una presencia significativa en nuestra vida cotidiana.

Desde tomar decisiones en función de incertidumbres hasta evaluar riesgos en diferentes contextos, las probabilidades nos brindan una herramienta poderosa para enfrentar la incertidumbre en una variedad de situaciones.

La aplicación de Business Process Reengineering (BPR) en una empresa supone un cambio radical en la forma de trabajar, Por supuesto, aquí tienes un grupo de ejercicios de teoría de la probabilidad junto con sus soluciones:

1. Probabilidades en el Clima

La predicción del clima es una aplicación ampliamente conocida de la teoría de probabilidades. Los pronósticos meteorológicos a menudo indican la probabilidad de lluvia, nieve o sol para un día determinado. Estos pronósticos se basan en datos históricos y modelos matemáticos que evalúan las probabilidades de diferentes condiciones climáticas.

Ejemplo:

Si el pronóstico del clima indica una probabilidad del 30% de lluvia, esto significa que existe un 30% de posibilidades de que llueva durante el día.

2. Probabilidades en Juegos de Azar

Los juegos de azar, como las loterías y los casinos, se basan en la teoría de probabilidades. Los participantes

apuestan en función de las probabilidades de ganar, que a menudo son muy bajas. Estas probabilidades están diseñadas para favorecer al establecimiento y generar ingresos.

Ejemplo:

En una lotería con un premio mayor de 1 millón de dólares y un costo de boleto de $1, la probabilidad de ganar es de 1 en 1,000,000. Esto significa que, en promedio, una persona debe gastar $1,000,000 para ganar $1,000,000.

3. Probabilidades en la Salud y Medicina

La teoría de probabilidades también se aplica en la medicina para evaluar riesgos y tomar decisiones clínicas. Las pruebas médicas, como las mamografías y los análisis de sangre, se basan en la probabilidad de detección de ciertas condiciones médicas.

Ejemplo:

Si una mamografía tiene una sensibilidad del 90% para detectar cáncer de mama y una especificidad del 95% para descartarlo, y la prevalencia de cáncer de mama en la población es del 1%, ¿cuál es la probabilidad de que una mujer con resultados positivos realmente tenga cáncer de mama?

P(Tiene cáncer | Resultado positivo) = (P(Resultado positivo | Tiene cáncer) * P(Tiene cáncer)) / P(Resultado

positivo)

P(Tiene cáncer | Resultado positivo) = (0.9 * 0.01) / (0.9 * 0.01 + 0.05 * 0.99) ≈ 0.15

4. Probabilidades en la Economía y Finanzas

Las probabilidades se utilizan en la toma de decisiones financieras y en la evaluación de riesgos en inversiones. El mercado de valores y las fluctuaciones de precios también se analizan utilizando métodos probabilísticos.

Ejemplo:

Si un inversor evalúa la probabilidad de que el precio de una acción aumente en un cierto período de tiempo en un 60%, puede tomar decisiones informadas sobre si invertir en esa acción.

La teoría de probabilidades es mucho más que una rama abstracta de las matemáticas; es una herramienta que se aplica en diversas áreas de la vida cotidiana. Desde el clima hasta los juegos de azar, desde la medicina hasta la toma de decisiones financieras, las probabilidades nos ayudan a evaluar incertidumbres y a tomar decisiones informadas.

Los ejemplos resueltos ilustran cómo estas aplicaciones se traducen en situaciones reales, mientras que las citas bibliográficas respaldan la importancia y la utilidad de la teoría de probabilidades en el mundo actual.

Ejercicios propuestos

Ejercicio 1

En una inversión, la probabilidad de obtener un rendimiento positivo es del 70%, mientras que la probabilidad de obtener un rendimiento negativo es del 30%. Si se invierte en dos activos independientes con estas mismas probabilidades de rendimiento, ¿cuál es la probabilidad de obtener al menos un rendimiento positivo?

Solución:

- P(Al menos un rendimiento positivo) = P(Rendimiento positivo en A) + P(Rendimiento positivo en B) - P(Rendimiento positivo en A y B)

- P(Al menos un rendimiento positivo) = 0.7 + 0.7 - (0.7 * 0.7) = 0.91

Ejercicio 2

En un ensayo clínico, la probabilidad de que un nuevo medicamento sea eficaz es del 80%. Si se administra el medicamento a 15 pacientes independientes, ¿cuál es la probabilidad de que al menos 10 de ellos experimenten mejoras?

Solución:

- P(Al menos 10 pacientes experimentan mejoras) = P(10 pacientes) + P(11 pacientes) + ... + P(15 pacientes)

- Utilizando la distribución binomial: P(X = k) = (nCk) * p^k * (1-p)^(n-k)

- P(Al menos 10 pacientes experimentan mejoras) = Σ[P(X=k)] desde k=10 hasta 15

Ejercicio 3

En un proceso de fabricación, la probabilidad de que una pieza sea defectuosa es del 10%. Si se seleccionan 5 piezas al azar, ¿cuál es la probabilidad de que al menos una de ellas sea defectuosa?

Solución:

- P(Al menos una pieza defectuosa) = 1 - P(Todas las piezas sean no defectuosas)

- P(Todas las piezas sean no defectuosas) = (0.9)^5

- P(Al menos una pieza defectuosa) = 1 - (0.9)^5

Ejercicio 4

En una encuesta sobre preferencias políticas, se sabe que el 60% de los encuestados apoya al Partido A. Si se seleccionan 8 encuestados al azar, ¿cuál es la probabilidad de que al menos 4 de ellos apoyen al Partido A?

Solución:

- P(Al menos 4 encuestados apoyen al Partido A) = P(4 encuestados) + P(5 encuestados) + ... + P(8 encuestados)

- Utilizando la distribución binomial: P(X = k) = (nCk)

* p^k * (1-p)^(n-k)

- P(Al menos 4 encuestados apoyen al Partido A) = Σ[P(X=k)] desde k=4 hasta 8

Ejercicio 5

En una empresa de distribución, la probabilidad de que un pedido sea entregado a tiempo es del 85%. Si se tienen 10 pedidos independientes, ¿cuál es la probabilidad de que al menos 8 de ellos sean entregados a tiempo?

Solución:

- P(Al menos 8 pedidos sean entregados a tiempo) = P(8 pedidos) + P(9 pedidos) + P(10 pedidos)

- Utilizando la distribución binomial: P(X = k) = (nCk) * p^k * (1-p)^(n-k)

- P(Al menos 8 pedidos sean entregados a tiempo) = Σ[P(X=k)] desde k=8 hasta 10

CAPÍTULO III

MODELOS DETERMINÍSTICOS Y PROBABILÍSTICOS

Existen aplicaciones en las que se dispone de un modelo que presenta una relación exacta entre las variables de interés; por ejemplo, la ley que describe el tiempo que tarda en caer un objeto desde una Altura dada o la fórmula que nos indica el interés ganado por un capital, datos la tasa de interés y el periodo de la inversión. Tales modelos se denominan deterministas.

Sin embargo, en la vida diaria, rara vez se presentan fenómenos que reproducen con exactitud una ley ya sea porque existen errores en la medición o porque hay otras que no son consideradas, por su escasa influencia, pero que son suficientes para que el modelo propuesto no sea exacto.

En consecuencia, un modelo en el que una o más variables es de naturaleza aleatoria se denomina probabilístico y a la determinación y examen de la calidad del modelo encontrado se llama análisis de regresión

Algunas de las más importantes aplicaciones del análisis de regresión son:

- Descripción cuantitativa de las relaciones entre una variable dada y un conjunto de variables.
- Interpolación entre valores de la función.
- Predicción y pronóstico de datos.

El interés será determinar una ecuación que relacione

una variable dada con otra variable de respuesta, bajo el supuesto que ellas se vinculan mediante una ecuación lineal de primer grado, caso particular conocido como regresión lineal simple.

1.1 Modelo lineal simple

Su ecuación de recta es:

$$y=\beta_0+\beta_1 x$$

Donde:

y: Variable endógena, explicada o respuesta

β_0: Interpretación de la recta con el eje y o intercepto

β_1: Pendiente de la recta

x: Variable exógena,explicativa o predictora

ε: Error aleatorio

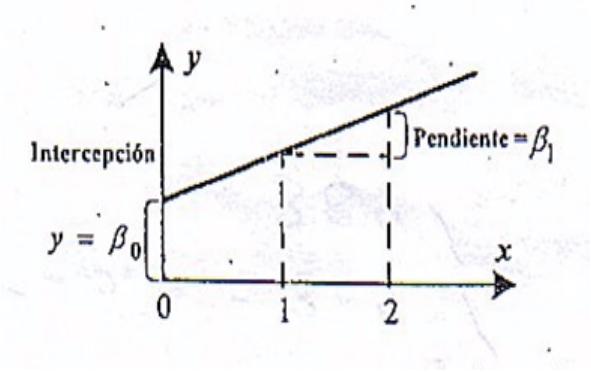


Figura 5. Ecuación del modelo lineal simple

Este modelo es determinista porque no considera el error y los valores e se obtienen, de manera exacta, al sustituir los valores de en la ecuación de la recta. Sin embargo, cuando se desea incorporar al modelo determinista el efecto aleatorio de las variables se le añade un componente que corresponde al error y el modelo queda como:

$$y=\beta\_(0\ )+\beta\_(1\ )\ x + \varepsilon$$

Para recoger el efecto aleatorio del error se plantean las siguientes hipótesis respecto a :

- Se distribuye normalmente con media cero y varianza .
- Errores correspondientes a observaciones distintas son independientes entre sí: .

Tabla 1. Ejemplos de modelos de regresión

Presupuestos de gastos de un hogar	Número de miembros del hogar	Efecto del nivel socioeconómico, tenencia de la vivienda, servicios que dispone, etc.
Precio de un departamento	Área de construcción	Efecto de la zona de ubicación, tipo de acabados, piso en el que se encuentra, etc.
Precio de un libro	Número de páginas del libro	Efecto del tipo de papel, la encuadernación, numero de ilustraciones, etc.

En análisis de regresión es necesario tener en cuenta

los siguientes pasos:

1) Tener una visión clara de objetivos del estudio con el fin de determinar cuál he de ser la varia respuesta y que variables puedan incluirse como variables independientes.

2) Recopilar los datos correspondientes a las variables identificadas como dependiente e independiente.

3) Postular un modelo al que se supone se ajustan los datos, en este caso se presume que es el lineal simple.

4) Determinar la ecuación de regresión; es decir, estimar los coeficientes del modelo propuesto.

5) Comprobar estadísticamente la adecuación del modelo. Incluye la realización de pruebas estadísticas sobre los parámetros, la ejecución de transformaciones de las variables para tener un mejor ajuste o retirar variables de una ecuación si su aporte no es significativo en la ejecución de predicción.

6) Cuando la ecuación sea satisfactoria, usar el modelo para realizar predicciones o estimaciones.

Una vez que se han cumplido los tres primeros pasos, su objetivo será estimar coeficientes del modelo y comprobar la adecuación del modelo.

3.1.1 Estimadores de parámetros

Se ocupa de deducir estimadores puntuales de parámetros β_1,β_2,σ^2 tal que se deducirá sus

principales propiedades usando únicamente hipótesis .

Teorema. La recta de regresión para por el punto $(\bar{x},\bar{y})$.

Se puede comprender mejor la fórmula $\mathrm{Cov}(\hat{\beta}_1,\hat{\beta}_2) = -\frac{\sigma^2\bar{x}}{S_{xx}}$ En la primera gráfica de la figura siguiente, es negativa pues cuando aumenta la pendiente de la recta, el punto de corte al eje aumenta proporcionalmente. En la gráfica derecha $\bar{x}$ es positiva tal que cuando aumenta la pendiente de la recta, el punto de corte al eje Y baja.

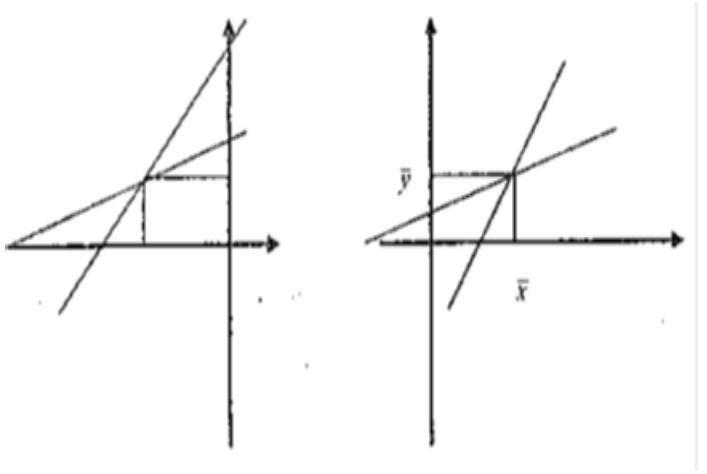


Figura 6. Comportamiento de las gráficas al aumentar la pendiente de la recta

Definición. Las diferencias $\hat{u}_i = y_i - \hat{y}_i$ se llaman residuos. Estos dan cuenta de errores no observables.

Observación. De ecuaciones $\begin{cases} \bar{y} - b_1 - b_2\bar{x} = 0 \\ \sum y_i x_i - n\bar{x}b_1 - b_2\sum x_i^2 = 0 \end{cases}$ se deduce:

$$\sum x_i \hat{u}_i = 0, \sum \hat{u}_i = 0$$

Tal que, la igualdad $\sum \hat{u}_i = 0$ puede ser falsa en un modelo que no incluye el término constante (intercepto) β_1.

Lema. En toda regresión lineal $\sum_{i=1}^{n} \hat{u}_i \hat{y}_i = 0$.

Notación. $SRC = \sum_{i=1}^{n} \hat{u}_i^2, SEC = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$ y $STC = \sum_{i=1}^{n} (y_i - \bar{y})^2$.

Donde:

SRC: Suma de Residuos al Cuadrado

SEC: Suma de Estimaciones al Cuadrado

STC: Suma Total de Cuadrado

Teorema. Si β_1 está en modelo o si la regresión no incluye el término constante, pero $\bar{y}=0$ tal que $STC_{\text{(Suma Total de Cuadrados)}} = SEC_{\text{(Suma de Estimaciones al Cuadrado)}} + SRC_{\text{(Suma de Residuos al Cuadrado)}}$

Demostración:

$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} [(y_i - \hat{y}_i) + (\hat{y}_i - \bar{y})]^2 = SRC + SEC + \sum_{i=1}^{n} (y_i - \hat{y}_i)(\hat{y}_i - \bar{y})$ tal que tal que $\sum_{i=1}^{n} (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) = \sum_{i=1}^{n} \hat{u}_i \hat{y}_i - \bar{y} \sum_{i=1}^{n} \hat{u}_i = 0$ la igualdad a cero se tiene por ecuaciones $\begin{cases} \bar{y} - b_1 - b_2 \bar{x} = 0 \\ \sum y_i x_i - n\bar{x} b_1 - b_2 \sum x_i^2 = 0 \end{cases}$ se deduce y lema $\sum \hat{u}_i \hat{y}_i = 0$.

Teorema. Bajo hipótesis H y si el término constante β_1 está en modelo tal que: $s_R^2 = \frac{SRC}{n-2}$

Es estimador insesgado de σ^2.

Demostración:

$y_i - \hat{y}_i = (y_i - \bar{y}) - \hat{\beta}_2 (x_i - \bar{x})$. Por $\hat{\beta}_2 = \frac{S_{xy}}{S_{xx}}, \hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x}$ tal que $y_i - \hat{y}_i = \beta_2 (x_i - \bar{x}) + (u_i - \bar{u}) - \hat{\beta}_2 (x_i - \bar{x}) = (u_i - \bar{u}) - (\hat{\beta}_2 - \beta_2)(x_i - \bar{x})$.

En consecuencia, $SRC = \sum_{i=1}^{n} (u_i - \bar{u})^2 - 2(\hat{\beta}_2 - \beta_2) \sum_{i=1}^{n} (u_i - \bar{u})(x_i - \bar{x}) + S_{xx} (\hat{\beta}_2 - \beta_2)^2$.

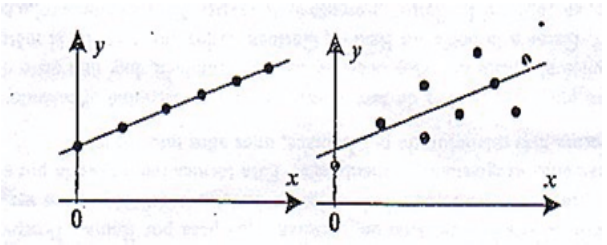
Las esperanzas de términos primero y tercero se calculan de manera sencilla. Viendo el segundo:

$$\hat{\beta}_2 = \frac{1}{S_{xx}} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{S_{xx}} \sum_{i=1}^{n} (x_i - \bar{x})[\beta_2 (x_i - \bar{x}) + (u_i - \bar{u})]$$

$$= \beta_2 + \frac{1}{S_{xx}} \sum_{i=1}^{n} (x_i - \bar{x})(u_i - \bar{u})$$

De donde:

$$\sum_{i=1}^{n} (x_i - \bar{x})(u_i - \bar{u}) = S_{xx} (\hat{\beta}_2 - \beta_2)$$



Reemplazando el segundo término:

$$SRC = \sum_{i=1}^{n} (u_i - \bar{u})^2 - S_{xx} (\hat{\beta}_2 - \beta_2)^2$$

$$SRC = E\left(\sum_{i=1}^{n} (u_i - \bar{u})^2 \right) - S_{xx} Var(\hat{\beta}_2)^2$$

La primera esperanza es conocida en estadística básica:

$$E(SRC) = (n-1)\sigma^2 - \sigma^2 = (n-2)\sigma^2$$

Definición. La raíz cuadrada de estimación de varianza se llama error estándar o error típico. El error estándar de $\hat{\beta}_1$ ($\hat{\beta}_2$) es:

$$ee(\hat{\beta}_1) = \sqrt{S_R^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}, ee(\hat{\beta}_2) = \sqrt{\left(\frac{S_R^2}{S_{xx}}\right)}$$

Tal que las fórmulas que continúan son útiles para estimar sumas de cuadrados.

Teorema. STC=S_{yy}. Si el término constante está en modelo o si datos están centrados:

$$SRC = S_{yy} - \left(\frac{S_{xy}^2}{S_{xx}}\right) \text{ i } SEC = \left(\frac{S_{xy}^2}{S_{xx}}\right)$$

Demostración. Por $y_i - \hat{y}_i = (y_i - \bar{y}) - \hat{\beta}_2(x_i - \bar{x})$ tal que $SRC = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}\left[(y_i - \bar{y}) - \hat{\beta}_2(x_i - \bar{x})\right]^2 = S_{yy} - 2\hat{\beta}_2 S_{xy} + \hat{\beta}_2^2 S_{xx} = S_{yy} - \left(\frac{S_{xy}^2}{S_{xx}}\right)$. La segunda igualdad es consecuencia de STC = SEC + SRC.

3.2 Método de mínimos cuadrados

El método de MCO es el más común en el análisis de regresión, específicamente por ser mucho más intuitivo y matemáticamente más sencillo que el método de máxima verosimilitud. Además, por lo general los dos métodos proporcionan resultados similares.

El método de mínimos cuadrados ordinarios se atribuye al matemático alemán Carl Friedrich Gauss (1777-1855), considerado el matemático más grande del siglo XIX, además de uno de los tres matemáticos más importantes de todos los tiempos (Arquímedes y Newton

son los otros dos).

Para estimar los coeficientes de la ecuación de regresión se empleará el método de los mínimos cuadrados, consistente en minimizar la suma de los cuadrados de los errores; esto es, que si se nota a la ecuación de predicción por:

$$\hat{y} = b_o + b_1 x$$

Donde b_0 y b_1 son estimadores de β_0 y β_1, respectivamente; ellos deben ser tales que la suma de los cuadrados de las diferencias entre los valores observados de la variable respuesta y su estimación por la ecuación de regresión sea mínima. Si se dispone de pares de observaciones de las variables independientes y dependientes (x_1;y_1), (x_2;y_2), (x_n;y_n) y si son los valores de las predicciones de y.

$$\hat{y} = b_o + b_1 x_i$$

Los residuos de la predicción (errores) se calculan por:

$$y_1 = y_1 - \hat{y}_i$$

Se busca valores de b_0 y b_1 que minimicen la suma de los cuadrados de la también llamada suma de los cuadrados de los residuos:

$$SCE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(y_1 - \hat{y}_i)^2$$

$$= \sum_{i=1}^{n}[y_1 - (b_0 + b_1 x_1)]^2$$

Derivando SCE con respecto a b_0 y b_1 , e igualando el

resultado a cero se obtiene la ecuación:

$$\frac{\partial(SCE)}{\partial b_o} = -2\sum_{i=1}^{n}(y_1 - b_0 - b_1 x_i) = 0,$$

$$\frac{\partial(SCE)}{\partial b_1} = -2\sum_{i=1}^{n} x_i(y_1 - b_0 - b_1 x_i) = 0,$$

Su solución es:

$$b_1 \frac{\sum_{i=1}^{n}(x_1 - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_1 - \bar{x})^2} = \frac{SC_{xy}}{SC_{xx}}$$

$$b_o = \bar{y} - b_1\bar{x},$$

Donde $\frac{\sum_{i=1}^{n} x_i}{n}$ y $\bar{y} = \frac{\sum_{i01}^{n} y_i}{n}$ Son los promedios de los valores de las variables independientes y dependientes. Una vez obtenidos los valores de y se los sustituye en la ecuación, de esta manera queda esta la recta de predicción por mínimos cuadrados:

$$\hat{y} = b_o + b_1 x,$$

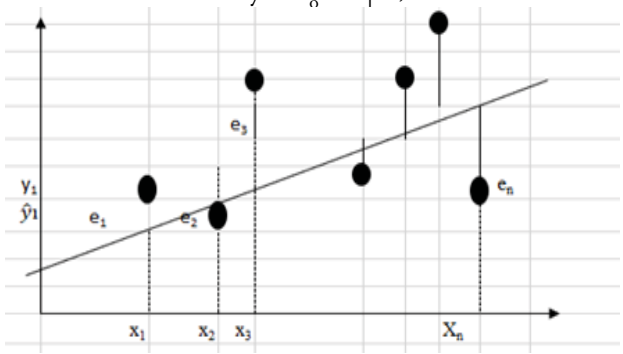


Figura 7. Comportamiento de las gráficas con base en los errores de nube de puntos

Observación. En la estimación de los parámetros se debe tener presente la incorporación de un redondeo en el cálculo de SC_{XY} y de SC_{Xx}de ; se recomienda el empleo de un número sufriente en cifras significativas al realizar los cálculos de forma manual.

Ejemplo:

En un estudio para determinar la relación entre el peso de los automóviles y su consumo de combustible se escogió una muestra de 10 carros, con resultados

Variable	Valores numéricos									
Consumo	8	16	6	7	7	9	11	12	1	20
Peso(kg)	739	1187	655	729	888	797	963	802	1551	1650

A partir de los siguientes supuestos el método de mínimos cuadrados presenta propiedades estadísticas muy atractivas que lo han convertido en uno de los más eficaces y populares del análisis de regresión, partiendo de la idea que el modelo de Gauss, modelo clásico o estándar de regresión lineal (MCRL) es el cimiento de la mayor parte de la teoría econométrica y plantea siete supuestos clásicos, en sentido que Gauss lo empleó por primera vez en 1821 y desde esta fecha sirve como norma o patrón con que compara modelos de regresión que no satisfacen los supuestos Gaussianos:

Modelo de Regresión es Lineal en los Parámetros.

Aunque, la variable regresada, dependiente o explicada (Y) y la regresora, independiente o explicativa (X) pueden o no ser lineal, inclusive puede incluir más variables explicativas, como se muestra enseguida:

$$Y_i=\beta_1+\beta_1 X_i+u_i$$

Valores fijos de , valores de independientes del término error. Los valores que toma la variable regresora, independiente o explicativa (X) pueden considerarse fijos en muestras repetidas (regresora fija, no aleatoria, Modelos de Mínimos Clásicos de Regresión Lineal –MCRL- o Regresora Fija) o haber sido muestreados junto con la variable regresada, dependiente o explicada (Y) (regresora estocástica, aleatoria o Modelos Noeclásico de Regresión Lineal –MNRL- o Regresora Estocástica). En el segundo caso, se supone que la(s) variable(s) X y término error son independientes; es decir, cov(X_i,u_i)=0.

El valor medio de perturbación u_i=0 (no hay error de especificación en modelo de regresión elegido). Dado el valor de X_i, la media o valor esperado del término de perturbación aleatoria u_i es cero. Simbólicamente es $E(u_i|X_i)=0 \Rightarrow E(Y_i|X_i)=\beta_1+\beta_1 X_i$ o, si X no es estocástica, equivale a $E(u_i)=0$, pues si la media condicional de una variable aleatoria, dada otra variable aleatoria, es cero, la covarianza entre las dos variables es cero y, por tanto, las dos variables no estás correlacionadas o X_i y u_i no están correlacionadas.

Cuando la Función de Regresión Poblacional

(FRP) se expresa en una ecuación, se supone que X y u, representando la influencia de todas las variables omitidas, ejercen influencias independientes y aditivas, en Y, pero si X y u están correlacionadas, no es posible evaluar los efectos de cada una sobre Y tal que si X y u tienen correlación positiva, X aumenta cuando u aumenta y, viceversa, disminuye cuando u disminuye. Asimismo, si X y u tienen correlación negativa, X se incrementa cuando u se reduce, disminuye cuando u aumenta.

Geométricamente, éste supuesto se representa:

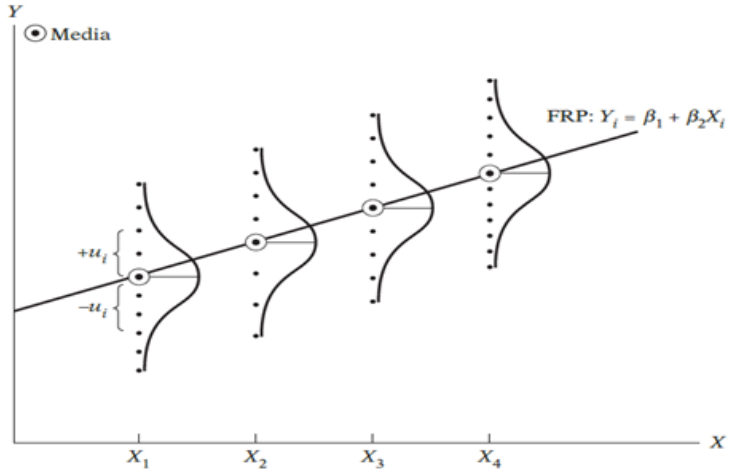


Figura 8. Representación geométrica del supuesto

Homocedasticidad o varianza constante de u_i (del griego skedanime que significa dispersar o esparcir tal que homo es igual y cedasticidad significa dispersión o, en otras palabras, igual varianza). La varianza del término error o de perturbación es la misma sin importan el valor de X.

Simbólicamente, se tiene:

$$\text{Var}(u_i) = E[u_i - (Y_i|X_i)]^2 = E(u_i^2|X_i) \text{ por supuesto 3}$$
$$= E(u_i^2) \text{si } X_i \text{ son variables no estocásticas} = \sigma^2$$

Esta ecuación establece que la varianza de u_i para cada X_i, varianza condicional de u_i, es algún número positivo constante igual a σ_2.

Por lo tanto, esta ecuación representa el supuesto de homocedasticidad. En términos llanos, la variación alrededor de la línea de regresión (relación promedio entre X y Y) es la misma para todos los valores de X tal que no aumente ni disminuye conforme varía X:

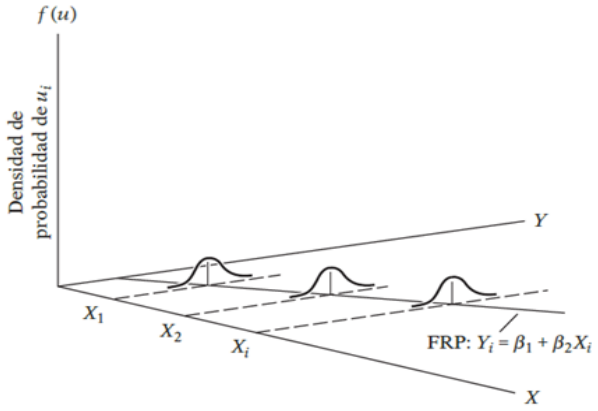


Figura 9. Variación alrededor de la línea de regresión (relación promedio entre X y Y) es la misma para todos los valores de X

Caso contrario, si se considera la siguiente figura

como varianza condicional de la población Y varia con X se conoce apropiadamente como heterocedasticidad o dispersión desigual o varianza desigual, escrita como $E(u_i^2 \mid X_i) = \sigma_i^2$ e indica con subíndice i que la varianza de la población Y no es constante:

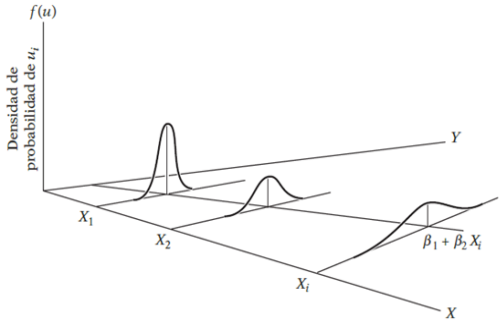


Figura 10. Varianza condicional de la población Y varia con X

No hay autocorrelación entre perturbaciones (perturbaciones u_i y u_j no están correlacionadas, supuesto de no correlación serial o no autocorrelación). Dados dos valores cualesquiera de X, X_i y X_j (i≠j), la correlación entre dos u_i y u_j=0 ∀(i≠j). En otras palabras, estas observaciones se muestran independientemente y simbólicamente es:

$$\text{cov}(u_i, u_j|X_i, X_j) = 0 \Rightarrow \text{cov}(u_i, u_j) = 0 \text{ si X no es estocástica}$$

Esto significa que, dado X_i, las desviaciones de dos valores cualesquiera de Y de sus valores promedio no muestran patrones como en figura a en que las u están correlacionadas positivamente, pues a una u positiva sigue una u igual o viceversa, mientras que en figura b

las u están correlacionadas negativamente, pues a una u positiva sigue una u negativa y viceversa:

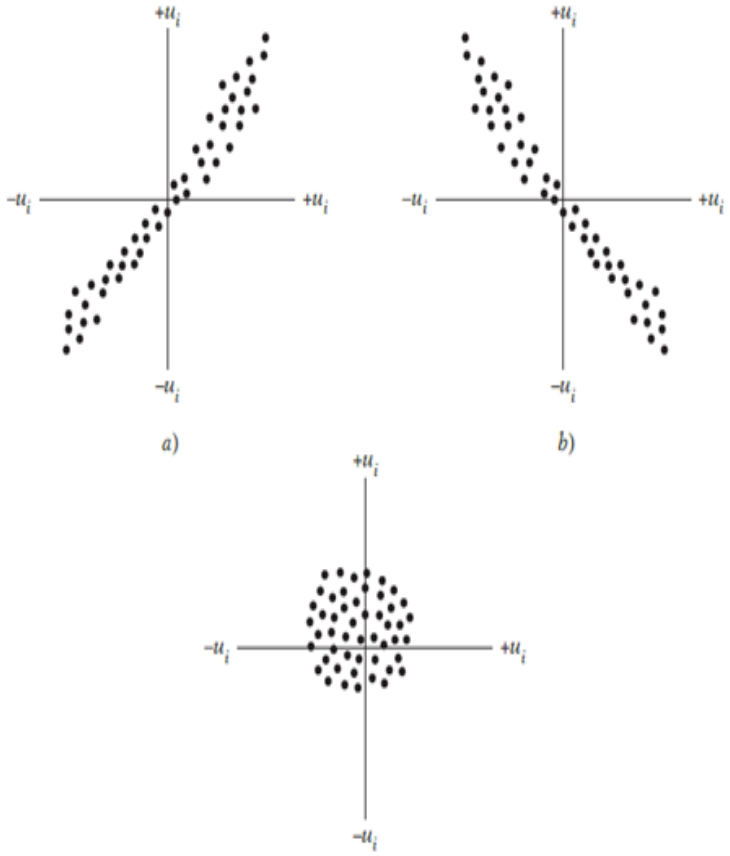


Figura 11. Patrones de las desviaciones de dos valores cualesquiera de Y respecto a sus valores promedio

Si las perturbaciones (desviaciones) siguen patrones sistemáticos, como figuras a y b, hay correlación serial o autocorrelación mientras que figura c muestra que no hay un patrón sistemático para las , indica cero

correlación.

- **Número de observaciones n debe ser mayor que número de parámetros por estimar. Sucesivamente, el número de observaciones será n≥número de variables explicativas.**

- **Naturaleza de variables X (variables deben variar). No todos los valores X en una muestra determinada deben ser iguales. Técnicamente, Var(X) debe ser un número positivo. Además, no puede haber valores atípicos de variable X; es decir, valores muy grandes en relación con el resto de las observaciones.**

Esto tiene base según ecuación $\hat{\beta}_2 = \frac{n\sum X_i Y_i - \sum X_i \sum Y_i}{n\sum X_i^2 - (\sum X_i)^2} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} = \frac{\sum x_i y_i}{\sum x_i^2}$, pues si todos los valores X son idénticos, $X_i = \bar{X}$ y el denominador de esta ecuación es cero (indeterminación) imposibilita la estimación de β_2 y, por consiguiente, de β_1. Por lo tanto, la variación tanto en Y como X es esencial para utilizar el análisis de regresión como herramienta de investigación.

Entonces, ¡las variables deben variar! El requisito que no existan valores atipícos de X es para evitar que resultados de regresión estén dominados por estos valores. Si hay algunos valores X que, por ejemplo, sean x veces el promedio de valores X, las líneas de regresión estimadas con o sin dichas observaciones serían muy diferentes. Con frecuencia estos valores atípicos son resultado de errores humanos de aritmética o de mezclar muestras de diferentes poblaciones.

Según (Castro, 2008), si se reemplaza parámetros desconocidos β_1,β_2 por dos números $b_1,b_2 \in R$ se produce una discrepancia entre $b_1 + b_2 x_i$ y y_i escrita como e_i tal que $e_i=y_i-(b_1 + b2\ x_i)$ donde la idea central es minimizar suma de desvíos al cuadrado:

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(y_i - b_1 - b_2 x_i)^2$$

Tal que se puede definir números $\hat{\beta}_1,\hat{\beta}_2$ que minimizan $SRC(b_1,b_2)=\sum_{i=1}^{n}(y_i-b_1- b_2x_i)^2$ donde $b_1,b_2 \in R$ se llaman estimadores de mínimos cuadrados (EMC) de parámetros β_1,β_2. Además, la notación:

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2, S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}), S_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2$$

Demostración:

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}(x_i - \bar{x})\,x_i = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$$

$$S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}$$

$$S_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(y_i - \bar{y})\,y_i = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2$$

Teorema. Bajo hipótesis H, estimadores de mínimos cuadrados (MC) de modelo $y_i=\beta_1+\beta_2 X_i+u_i$ i=1,2,3,4,…,n son:

$$\hat{\beta}_2 = \frac{S_{xy}}{S_{xx}} \text{ tq } \hat{\beta}_1 = \bar{y} - \hat{\beta}_2\bar{x}$$

Demostración. Para minimizar la función SRC $(b_1,b_2)=\sum_{i=1}^{n}(y^i-b^1- b^2 x^i)^2$ se estiman derivadas parciales:

$$\frac{\partial SCR(b_1,b_2)}{\partial(b_1)} = -2\sum_{i=1}^{n}(y_i - b_1 - b_2 x_i)$$

$$\frac{\partial SCR(b_1,b_2)}{\partial(b_2)} = -2\sum_{i=1}^{n}(y_i - b_1 - b_2 x_i)\,x_i$$

Igualando a cero, se tiene el sistema de ecuaciones llamado Sistema Normal de Ecuaciones:

$$\begin{cases} \bar{y} - b_1 - b_2\bar{x} = 0 \\ \sum y_i x_i - n\bar{x}\,b_1 - b_2\sum x_i^2 = 0 \end{cases}$$

Su solución está dada por el enunciado del teorema anterior.

3.2.1 Propiedades de estimadores

Definición. Cualquier estimador que sea combinación lineal de observaciones y_i se dice que es un estimador lineal, pues:

$$\hat{\beta}_2 = \frac{S_{xy}}{S_{xx}} = \frac{1}{S_{xx}}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{S_{xx}}\sum_{i=1}^{n}(x_i - \bar{x})y_i = \sum_{i=1}^{n} a_i y_i \text{ tal que } a_i = \frac{x_i - \bar{x}}{S_{xx}}$$

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2\bar{x} = \frac{1}{n}\sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\bar{x}a_i y_i = \sum_{i=1}^{n} b_i y_i \text{ tal que } b_i = \frac{1}{n} - \frac{(x_i - \bar{x})\bar{x}}{S_{xx}}$$

Teorema. Bajo hipótesis H, estimadores de MC son insesgados y:

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{S_{xx}}, Var(\hat{\beta}_1) = \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right), Cov(\hat{\beta}_1,\hat{\beta}_2) = -\sigma^2\left(\frac{\bar{x}}{S_{xx}}\right)$$

Demostración. Las igualdades $\hat{\beta}_1 = \bar{y} - \hat{\beta}_2\bar{x} = \frac{1}{n}\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \bar{x}a_i y_i$ y $\hat{\beta}_2 = \frac{S_{xy}}{S_{xx}} = \frac{1}{S_{xx}}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$ y facilitan estas demostraciones, tal que por linealidad de esperanza:

$$E(\hat{\beta}_2) = \sum_{i=1}^{n} a_i E(y_i) = \sum_{i=1}^{n} \frac{x_i - \bar{x}}{S_{xx}}(\beta_1 + \beta_2 x_i) = \beta_2 \sum_{i=1}^{n} \frac{x_i - \bar{x}}{S_{xx}}(x_i) = \beta_2$$

Tal que por correlación de variables y_i:

$$Var(\hat{\beta}_2) = \sum_{i=1}^{n} a_i^2 Var(y_i) = \sigma^2 \sum_{i=1}^{n} a_i^2 = \frac{\sigma^2}{S_{xx}}$$

Las otras igualdades se demuestran de forma semejante. La igualdad $Cov(\hat{\beta}_1, \hat{\beta}_2) = -\sigma^2\left(\frac{\bar{x}}{S_{xx}}\right)$ manifiesta que estimadores $\hat{\beta}_1, \hat{\beta}_2$ son no correlacionados sólo si $\bar{x}=0$, , caso contrario si $\bar{x}>0$, varia en sentidos opuestos; si la pendiente aumenta el punto de corte al eje Y baja:

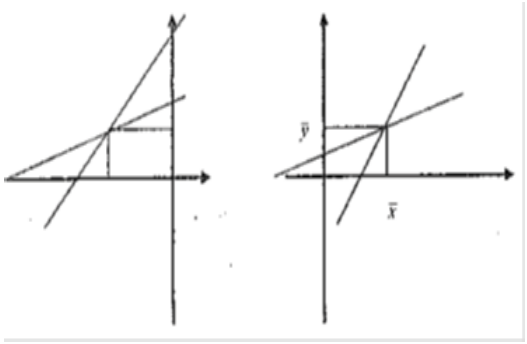


Figura 12. Patrones de desigualdades según aumento de la pendiente

Tal que se puede estimar la varianza de cualquier combinación lineal de estimadores:

$$Var(a\hat{\beta}_1 + b\hat{\beta}_2) = a^2 Var(\hat{\beta}_1) + 2abCovVar(\hat{\beta}_1, \hat{\beta}_2) + b^2 Var(\hat{\beta}_2)$$

$$= \sigma^2\left[a^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right) - 2ab\left(\frac{\bar{x}}{S_{xx}}\right) + \frac{b^2}{S_{xx}}\right]$$

$$= \sigma^2\left[\frac{a^2}{n} + \frac{1}{S_{xx}}(a^2\bar{x}^2 - 2ab\bar{x} + b^2)\right] = \sigma^2\left[\frac{a^2}{n} + \frac{(b - a\bar{x})^2}{S_{xx}}\right]$$

3.2.2 Propiedades de estimadores de mínimos cuadrados ordinarios según supuestos de normalidad

Suponga que u_i sigue una distribución normal, pues supone que cada u_i está Normalmente Distribuida con Media ($E(u_i)=0$), Varianza ($E[u_i-E(u_i)]^2=E(u_i^2)=\sigma^2$), $Cov(u_i,u_j)$ ($E\{[u_i-E(u_i)][u_j-E(u_j)]\}=E(u_i,u_j)=0$ $i\neq j$), expresados en forma resuma como $u_i \sim N(0,\sigma^2)$, por lo que con el supuesto de normalidad la ecuación anterior indica que u_i y u_j no están correlacionadas, sino que están distribuidas independientemente tal que u_i~Normal e Independientemente Distribuido $(0,\sigma^2)$, los estimadores de MCO tienen las propiedades siguientes:

1. Son Insesgados.

2. Tienen Varianza Mínima. En combinación con 1 significa que son estimadores insesgados con varianza mínima o eficientes.

3. Presentan Consistencia. A medida que el tamaño muestral aumenta indefinidamente, los estimadores convergen hacia sus verdaderos valores poblacionales.

4. $\hat{\beta}_1$ está Normalmente Distribuida al ser función lineal de u_i con Media ($E(\beta_1)=\beta_1$), Var, $\left((\hat{\beta}_1) \Rightarrow \sigma^2_{\hat{\beta}_1} = \frac{\sum x_i^2}{n\sum x_i^2}\sigma^2\right)$ en forma más completa $\hat{\beta}_1 \sim N(\beta_1, \sigma^2_{\hat{\beta}_1})$. Tal que, de acuerdo con las propiedades de distribución normal, variable Z es definida como $z = \frac{\hat{\beta}_1 - \beta_1}{\sigma_{\beta_1}}$ sigues una distribución normal estándar; es decir, una distribución normal con media cero y varianza unitaria (=1) ó, en otras palabras, Z~N(0,1).

5. $\hat{\beta}_2$ está Normalmente Distribuida al ser función lineal de u_i con Media ($E(\beta_2)=\beta_2$), Var, $\left((\hat{\beta}_2) \Rightarrow \sigma^2_{\hat{\beta}_2} = \frac{\sigma^2}{\sum x_i^2}\right)$ en forma más completa $\hat{\beta}_2 \sim N(\beta_2, \sigma^2_{\hat{\beta}_2})$

Entonces, $Z = \frac{\hat{\beta}_2 - \beta_2}{\sigma_{\beta_2}}$ sigue una distribución normal estándar.

Geométricamente, las distribuciones de probabilidad de $\hat{\beta}_1$ y $\hat{\beta}_2$ son:

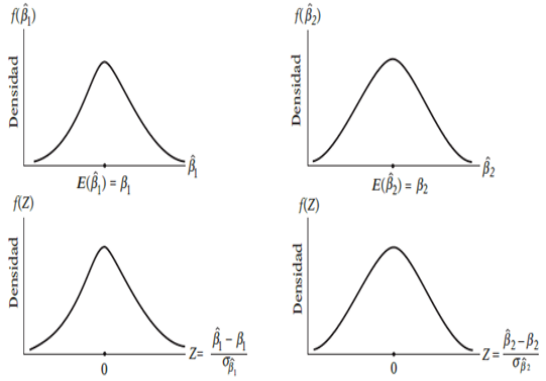


Figura 13. Distribuciones de probabilidad de y

6. Como la Distribución Ji Cuadrada (X^2) está

distribuida (n-2) $_{\text{(Grads de libertad)}}$ ($\hat{\sigma}^2/\sigma^2$). Esto ayuda a hacer inferencias respecto a la verdadera σ^2 a partir de $\hat{\sigma}^2$.

7.($\hat{\beta}_1$,$\hat{\beta}_2$) se distribuyen de manera independiente respecto a $\hat{\sigma}^2$.

8. $\hat{\beta}^1$ y $\hat{\beta}^2$ tienen varianza mínima entre todas las clases de estimadores insesgados, lineales o no lineales. Es eficaz debido a que, a diferencia del Teorema de Gauss-Markov, no se limita a la clase de estimadores lineales. Por tanto, los Estimadores de Mínimos Cuadrados son los Mejores Estimadores Lineales Insesgados (MELI), pues tienen varianza mínima en toda clase de estimadores insesgados.

Finalmente, el supuesto de normalidad permite derivar las distribuciones de probabilidad o muestrales de $\hat{\beta}_1$ y $\hat{\beta}_2$ (ambas normales) y de $\hat{\sigma}^2$ relacionadas con Ji Cuadrada (X^2).

CAPÍTULO IV

PRUEBAS DE HIPÓTESIS

Las pruebas de hipótesis en la regresión lineal múltiple se emplean para determinar la significación de la regresión; es decir, si globalmente las variables aportan información al modelo y realizar pruebas sobre valores de coeficientes individuales para examinar si una variable particular es significativa en modelo y merece ser incluida en la ecuación.

4.1 Intervalo de confianza

El intervalo de confianza delimita la región en que probablemente se encuentra el verdadero valor del parámetro tal que existe a lo sumo una probabilidad $\alpha=1-\eta$ que esté fuera del intervalo. Esto se usa para contrastar o probar las hipótesis:

$$\begin{cases} \mathbf{H_o}: \beta_1 = 0 \\ \mathbf{H_o} \text{ o } \mathbf{H_1}: \beta_1 \neq 0 \end{cases} \quad \begin{cases} \mathbf{H_o}: \beta_2 = 0 \\ \mathbf{H_o} \text{ o } \mathbf{H_1}: \beta_2 \neq 0 \end{cases}$$

H_0 se lee "H subíndice 0", pues H significa hipótesis y el subíndice 0 señala "no hay diferencia". En cambio, H_1 o H_a, leído como "H subíndice 1 o a" señala que "al menos 1 es diferente respecto al resto de los tratamientos".

La demostración matemática que lo sustenta se encuentra en la siguiente nota al pie y da razón que exista H_0 e H_a en una prueba de dos colas.

Además, con base en una muestra de una variable aleatoria X de ley P_θ en decidir entre dos hipótesis: $H_0: \theta \in \theta_{0\,(\text{hipótesis nula o privilegiada})}$ vs $H_1: \theta \in \theta_{1\,(\text{hipótesis alternativa o alterna})}$ tal que $\theta_0 \cup \theta_1 = \theta$ y $\theta_0 \cap \theta_1 = \varnothing$. Si d_0 y d_1 representan, respectivamente, las decisiones de no rechazar H_0 o H_1 y $D=\{d_0,d_1\}$.

Entonces: sea $X:\Omega \rightarrow E \subseteq R$ una variable aleatoria de ley P_θ. Se llama "Prueba de Hipótesis Pura" o test puro a toda aplicación: $\varnothing$: $E^n \rightarrow D$ tal que $\varnothing$ equivale a particionar E^n en dos conjuntos: $W=\varnothing^{-1}(d_0)=\{x \in E^n$:si se observa x,no se acepta H_0 } y W^c=$\varnothing$^(-1) $(d_1)=\{x \in E$^n:si se observa x,no se rechaza $H_0\}$. El nombre de H_0 proviene que H_0 se asumirá como verdadera, salvo que datos muestrales indiquen su falsedad. Nula debe entenderse como neutra.

H_0 nunca se considera probada o demostrada, salvo estudiando todos los datos de la población, puede diferir en un valor pequeño imperceptible para el muestreo, que puede ser imposible; aunque, puede no ser aceptada por los datos. Con base en esto, no se debe afirmar "se acepta H_0", siendo lo correcto "no se rechaza H_0" y, por abuso de lenguaje, es común hallar la expresión "se acepta la H_0" en lugar de "no se rechaza H_0".

Generalmente, H_0 se elige según con el principio de simplicidad científica, que establece que únicamente se abandona un modelo simple a favor de otro más complejo cuando la evidencia a favor de este último sea fuerte. Por el carácter dicotómico, si no se acepta H_0, automáticamente no se rechaza H_1.

Finalmente, se llama riesgo de primera especie a la probabilidad de rechazar H_0 cuando es verdadera: $\alpha_0(\varnothing)=P_0(W_{\text{(Región critica de prueba)}})$, mientras que el riesgo de segunda especie es la probabilidad de aceptar H_0 cuando es falsa: $\beta\theta\ (\varnothing)=P_1\ (W_{-\text{(Región de aceptacion)}}$^c)=1-$P_1(W)$. Se

llama "Potencia de una prueba" a la probabilidad de no aceptar H_0 cuando es falsa -P_1 $(W)=1-\beta\theta\ (\varnothing)$- y, también, se denomina "Nivel de significación de una prueba" al valor $\alpha=\sup \alpha_\theta\ (\varnothing)\ \ \theta \in \theta_0$.

Cuando $\theta \subseteq R$ se estudian problemas del tipo:

$$P_0:\begin{cases} H_0:\theta=\theta_0, \\ H_1:\theta=\theta_1 \text{ con } \theta_0 \neq \theta_1' \end{cases} \qquad P_1:\begin{cases} H_0:\theta \leq \theta_0, \\ H_1:\theta > \theta_{0\ \text{(Prueba de cola derecha)}}' \end{cases}$$

$$P_2:\begin{cases} H_0:\theta \geq \theta_0, \\ H_1:\theta < \theta_{0\ \text{(Prueba de cola izquierda)}}' \end{cases}$$

$$P_3:\begin{cases} H_0:\theta_0 \leq \theta \leq \theta_1, \\ H_1:\theta < \theta_{0\ \text{(Prueba de cola izquierda)}}\ o\ \theta > \theta_{1\ \text{(Prueba de cola derecha)}}(\theta_0 < \theta_1)' \end{cases}$$

$$P_4:\begin{cases} H_0:\theta < \theta_0\ o\ \theta > \theta_1, \\ H_1:\theta_0 \leq \theta \leq \theta_1, \end{cases} \text{ o } P_5:\begin{cases} H_0:\theta=\theta_0, \\ H_1:\theta \neq \theta_1. \end{cases}$$

4.2 Teorema Neyman-Pearson

El teorema Neyman-Pearson indica, en contexto de prueba propuesta: $\forall\ \alpha \in [0,1]$,$\epsilon$ un test puro $\varnothing$, de nivel α, de potencia máxima, definido por la región critica: $W=\{x \in E^n$: $L(\theta_0,x)/L(\theta_1,x) \leq k\} \longrightarrow$ se determina de la condición $\alpha=P_0(W)$. representa la función de verosimilitud y una observación de la muestra. Otras pruebas importantes, de nivel , son:

$$A):\begin{cases} H_0:\mu > \mu_0, \\ H_1:\ \mu \leq \mu_0 \end{cases} \text{ y } B):\begin{cases} H_0:\mu < \mu_0, \\ H_1:\ \mu \geq \mu_0 \end{cases}$$

Se define el estadístico razón de t de Student:

$$t_j = \frac{\hat{\beta}_j}{ee(\hat{\beta}_j)}$$

Corolario. Bajo hipótesis N, si el término constante está en regresión o si datos han sido centrados. Se rechaza hipótesis nula (H_o) a nivel α si sólo si:

$$|t_j| \geq t_{n-2}\left(\frac{\alpha}{2}\right)$$

Tal que, el nivel α de una prueba es una cota superior para la probabilidad de cometer un error tipo I, rechazar H_o cuando es verdadera. Cuando se hace una prueba es recomendable estimar la probabilidad crítica (ρ-Value), que es valor mayor de α en que no se rechaza H_o.

Es decir, ρ-value es la probabilidad de observar un valor muestral tan extremo o más extremo que el valor observado dado que la H_0 es verdadera o es una manera de expresar la probabilidad que H_0 no sea verdadera o, según (Galindo, 2006), "mínimo valor del nivel de significación para que datos observados indican que H_0 será rechazada".

Complementariamente, Pr(>F) es el nivel de probabilidad que H_a caiga en la zona de rechazo. Las investigaciones sociales usualmente trabajan con α=0.10 o 90 % de confiabilidad estadística, Diseños Experimentales con α≤0.05 o 95 % de confiabilidad estadística, Control de Calidad α≤0.01 o 99 % de confiabilidad estadística y Control de Calidad con 6σ α≤0.001 o ≥99.9 % de confiabilidad estadística.

Demostración:

No se rechaza H_o si i sólo si:

$$|t_j| < t_{n-2}\left(\frac{\widehat{\alpha}}{2}\right)$$

Como la función de distribución es monótona creciente tal que:

$$\Pr(T_{n-2} < |t_j|) = 1 - \frac{\widehat{\alpha}}{2} \text{ tal que } \widehat{\alpha} = 2\Pr(T_{n-2} \geq |t_j|)$$

Además, estimar la potencia de prueba es importante, pues la probabilidad de rechazar H_o cuando H_a o H_1 es la hipótesis verdadera; es decir, es correcto rechazar H_o. Cuando $B_j=b_j\neq 0$ si se llama η a función potencia:

$$\eta(b_j) = \Pr\left(|t_j| < t_{n-2}\left(\frac{\widehat{\alpha}}{2}\right) \middle| B_j = b_j\right)$$

Bajo hipótesis $B_j=b_j$, la razón t_j sigue una ley de Student descentrada con n-2 grados de libertad y parámetro descentramiento $\delta^2=(bj^2)/Var(B^j)$,que puede ser aproximada por una ley de Student centrada. Los paquetes estadísticos presentan potencias de pruebas para $B_j=B_j$ y reemplazando $Var(B_j)$ por $(Var)^{\wedge}(B_j)$; es decir, se reemplaza σ^2 por S_R^2. Si rechaza que pendiente de recta es nula se interpreta como regresión es significativa.

Para probar la significancia de regresión se acostumbra usar prueba de Fisher, que en caso de regresión lineal simple es equivalente a prueba t de Student para la pendiente. El cálculo del estadístico F de Fisher o F de Snedecor se acostumbra escribir la tabla de análisis de varianza (ANOVA,ADEVA,ANDEVA,ANVA,AVAR o

ANOVA de Fisher).

Una tabla ANOVA descompone la varianza es diferentes rubros, que en regresión lineal se descompone en varianza explicada por regresión y varianza debido al error.

En otras palabras, está asociado a un nivel crítico de probabilidad de obtener valores, como el obtenido o mayores. Si éste es mayor al nivel de error asumido, habitualmente 95 % de confiabilidad estadística, error o nivel de significancia (α) igual a 0.05, se interpreta "con base en resultados obtenidos de datos muestrales (poblacionales), no se acepta hipótesis nula, que indica igualdad de medias de tratamientos (H_0: $\mu_1=\mu_2=\mu_3=\mu_4=\cdots=\mu(k)=\mu$ (0) o H0: $\tau_1=\tau_2=\tau_3=\tau_4=\cdots\tau_k=0$) y no se rechaza la alternativa, que indica diferencia de al menos una media de tratamiento respecto al resto (Ha: $\mu_i\neq\mu_j$ para algún $i\neq j$ o H_a: $\tau_i\neq 0$ para algún i)".

Con diferencia parcial respecto a (Gujarati & Porter, 2010) que afirman "con base en una prueba de significancia, como prueba t, se dice "aceptar" la hipótesis nula, todo lo que se afirma es que, con base en la evidencia dada por la muestra, no existe razón para rechazarla, no se sostiene que la hipótesis nula sea verdadera con absoluta certeza" y, de acuerdo con (González, 1985), "de la misma manera que un tribunal se pronuncia un veredicto de "no culpable" en vez de "inocente", así la conclusión de una prueba estadística

es **"no rechazar" en vez de "aceptar".** $|F\_Calculada| > |F_{-Tablas\ (\alpha,\ k-1,N-k)}|$ F_{Tablas} de Fisher Tablas ($\alpha_{(Tamaño\ de\ error\ o\ nivel\ de\ significancia)}$, $k-1_{(Grados\ de\ libertad\ del\ numerador)}$, $N-k_{-(Grados\ de\ libertad\ del\ denominador)}$, se tiene el criterio más formal y/o riguroso de lectura de una ANOVA que no se acepta hipótesis nula, que indica igualdad de medias de tratamientos (H0: $\mu_1=\mu_2=\mu_3=\mu_4=\cdots=\mu_k=\mu$ (0) o H_0: $\tau_1=\tau_2=\tau_3=\tau_4=\cdots\tau_k=0$) y no se rechaza la alternativa, que indica diferencia de al menos una media de tratamiento respecto al resto (H_a: $\mu_i\neq\mu_j$ para algún $i\neq j$ o H_a: $\tau_i\neq 0$ para algún i).

Si

$|F_{Calculada}| < |F_{(Tablas(\alpha,k-1,N-k)}|$ F_{Tablas} de Fisher Tablas ($\alpha_{(Tamaño\ de\ error\ o\ nivel\ de\ significancia)}$, $k-1_{(Grados\ de\ libertad\ del\ numerador)}$, $N-k_{(Grados\ de\ libertad\ del\ denominador)}$) se tiene el criterio más formal y/o riguroso de lectura de una ANOVA que no se rechaza hipótesis nula, que indica igualdad de medias de tratamientos (H_0: $\mu_1=\mu_2=\mu_3=\mu_4=\cdots=\mu(k)=\mu$ (0) o H_0: $\tau_1=\tau_2=\tau_3=\tau_4=\cdots\tau_k=0$) y no se acepta la alternativa, que indica diferencia de al menos una media de tratamiento respecto al resto (Ha: $\mu_i\neq\mu_j$ para algún $i\neq j$ o Ha: $\tau_i\neq 0$ para algún i).

Tabla **2**. *Análisis de varianza de Fisher*

Análisis de Varianza o ANOVA,ADEVA,ANDEVA,ANVA,AVAR o ANOVA de Fisher					
F. de V(Factor de Varia o VF(Variation factor)	GI(Grados de Libertad) o Df(Degrees of freedom)	SC(Suma de Cuadrados) o Sum Sq(Sum of square)	CM(Cuadrado Medio) o Mean Sq(Mean squares)	F(calculada) o F value(Calculated F)	Pr(> F) o ρ – Value
Total	(n – 1)	STC(Suma Total de Cuadrados)			
Intergrupo o Regresión	(1)(gl numerador)	SEC(Suma de Estimaciones al Cuadrado)	SEC/gl(numerador)	$\frac{C.M_{Regresión}}{C.M_{Error}}$	$\frac{gl_{(numerador)}}{gl_{(denominador)}}$
Intragrupo o Error	(n – 2)(gl denominador	SRC(Suma de Residuos al Cuadrado)	SRC/gl(denominador)		

Las sumas de cuadrados en que la regresión tiene término constante o datos han sido centradas. Las medias o promedios de sumas de cuadrados se obtienen dividiendo las sumas de cuadrados por grados de libertad. Es decir, son el número de contrastes ortogonales menos el número de restricciones impuestas, que pueden hacerse en un grupo de datos.

Los grados de libertad pueden descomponerse al igual que la suma de cuadrados en Gl total = Gl entre grupos o Tratamientos + Gl dentro de grupos o Error; aunque, sus divisiones pueden aumentar según el diseño experimental que se trabajó.

Además, refiere a un número de valores a escoger libremente son el número de datos que son libres de variar cuando se calcula tal prueba o indican el número de dimensiones en que vectores pueden variar libremente, introducida la expresión por Fisher, de un conjunto de observaciones están dados por el número de valores que pueden ser asignados arbitrariamente, antes que el resto de variables queden completamente determinadas.

Según (Castro, 2008), los grados de libertad corresponden a dimensiones de sub espacios vectoriales en que se encuentran los datos tal que un conjunto generador se representa a partir que vectores $\upsilon_1,\upsilon_2,\upsilon_3,\upsilon_4,\ldots,\upsilon n$ de un espacio vectorial V genera a V si todo vector en V se puede escribir como una combinación lineal de los mismos; es decir, $\forall\ \upsilon \in V_{(\text{Espacio Vectorial})}$) ε escalares

$\alpha_1,\alpha_2,\alpha_3,\alpha_4,\ldots,\alpha n \Rightarrow \upsilon=\alpha_1\ \upsilon_1+\alpha_2\ \upsilon_2+\alpha_3\ \upsilon_3+\alpha_4\ \upsilon_4\ldots+\alpha_n\ \upsilon_n$.

Además, un espacio generado por un conjunto de vectores se genera a partir que sea υ1,υ2,υ3,υ4,…,υk k vectores de un espacio vectorial V. El espacio generado por$\{\upsilon_1,\upsilon_2,\upsilon_3,\upsilon_4,\ldots,\upsilon k\}$ es el conjunto de combinaciones lineales $\upsilon_1,\upsilon_2,\upsilon_3,\upsilon_4,\ldots,\upsilon_k$; es decir, gen$\{\upsilon_1,\upsilon_2,\upsilon_3,\upsilon_4,\ldots,\upsilon k\}=\{\upsilon \mid \upsilon=\alpha_1\ \upsilon_1+\alpha_2\ \upsilon_2+\alpha_3\ \upsilon_3+\alpha_4\ \upsilon_4\ldots+\alpha_k\ \upsilon_k\}$.

Por lo tanto, se dice que es linealmente dependiente (L.D.o D.L.en inglés) si sean $\alpha 1\ \upsilon_1+\alpha_2\upsilon_2+\alpha_3\ \upsilon_3+\alpha_4\ \upsilon_4\ldots+\alpha_n\ \upsilon_n$, n vectores en espacio vectorial V tal que los vectores son linealmente dependientes si existen n escalares $\alpha_1,\alpha_2,\alpha_3,\alpha_4,\ldots,\alpha_n$ NO TODOS SON IGUALES A CERO; es decir, con algún $\alpha_i\neq 0 \Rightarrow \alpha_1\ \upsilon_1+\alpha_2\upsilon_2+\alpha_3\ \upsilon_3+\alpha_4\upsilon_4\ldots+\alpha n\ \upsilon n=0\vec{}$.

Caso contrario, estos serán linealmente independientes (L.I.o I.L.en inglés) tal que, la única combinación lineal que da cero en la TRIVIAL es $\alpha_1=\alpha_2=\alpha_3=\alpha_4=\cdots=\alpha_n=0$. Por ejemplo: $STC_{(\text{Suma Total de Cuadrados})}$ es norma del vector cuyas componentes son $\{y_i-\bar{y} \mid i=1,2,3,4,\ldots,n\}$ que se encuentra en un sub espacio vectorial de R^n de dimensión n-1. $SEC_{(\text{Suma de Estimaciones al Cuadrado})}$ es suma de predicciones que se hallan sobre una recta tal que está asociada a un sub espacio vectorial de dimensión 1.

Para la suma $\sum^n_{i=1} y^2_i$ los grados de libertad son n, pues es la norma del vector $(y_1,y_2,y_3,y_4,\ldots,y_n)\epsilon$ R tal que para $SRC_{(\text{Suma de Residuos al Cuadrado})}$ en una regresión sin término constante los grados de libertad son n-1.Sin embargo, si los datos han sido centrados $\sum y_i=0$, el vector $(y_1,y_2,y_3,y_4,\ldots$

$,y_n)$ se encuentra sobre un sub espacio vectorial de R^n de dimensión n-1.

Teorema.

Bajo hipótesis N, si la regresión tiene término constante para $\beta_2=0$ tal que la razón F sigue la ley Fisher-Snedecor con (1,n-2) grados de libertad $(F\longrightarrow F_{(1,n-2)})$. Este teorema permite no aceptar H_0: $\beta_2=0$ a favor de H_a o H_1: $\beta_2\neq 0$ al nivel de α si $F_{(\text{Calculado})}$es superior al fractil $(F_{(1,n-2)},\alpha)$ de ley de Fisher con (1,n-2) grados de libertad, de orden 1-α; es decir, la cola derecha de la función de densidad es igual a α.

La probabilidad crítica para esta prueba es:

$$\hat{\alpha} = \Pr\left(F_{(1,n-2)} \geq F_{(\text{Tablas})}\right)$$

De acuerdo con (González, 1985), "de la misma manera que un tribunal se pronuncia un veredicto de "no culpable" en vez de "inocente", así la conclusión de una prueba estadística es **"no rechazar" en vez de "aceptar"**. Esto se basa, en parte, en las siguientes razones:

- (Wackerly, Mendehall, & Scheaffer, 2010) señalan "la probabilidad recibe el nombre de nivel de significancia o, en forma más sencilla, nivel de prueba. Aun cuando se recomiendan con frecuencia pequeños valores de , el valor real para usar en un análisis es un tanto arbitrario. Sin embargo, es el nivel de

significancia alcanzado, relacionado con una prueba y es un estadístico que representa el valor más pequeño de para el cual la información muestral indica que puede ser rechazada. En cierto sentido, permite al lector de la investigación evaluar la magnitud de la discrepancia entre datos observados e H_0"..

• (Navidi, 2006) afirma "si se rechaza H_0 se concluye que era falsa, en cambio si H_0 no se rechaza no se concluye que H_0 es verdadera, pues sólo se puede concluir que H_0 es factible, pero nunca se puede llegar a la conclusión H_0 es verdadera debido a que, por un lado, el estadístico de prueba es consistente con hipótesis H_a y está un poco en desacuerdo con H_0 tal que la única cuestión es si el nivel de desacuerdo medido con ρ-value es suficientemente grande para presentar H_0 como no factible y, por otro lado, una regla general conveniente, considerando ρ-value como menor que un umbral específico, indica rechazar o no aceptar H_0 cada vez que $\rho \leq 0.05$, interpretando que es "estadísticamente significativo"; aunque, este criterio no tiene ninguna base científica y, además, ésta es una mala práctica por varias razones:

• No proporciona ninguna manera de decir si ρ-value era apenas menor que 0.05 o si era mucho menor,

• Notificar que un resultado era estadísticamente significativo a un nivel de 5% implica que hay gran diferencia entre ρ-value justo debajo de 0.05 y uno justo arriba de 0.05, cuando efectivamente hay una diferencia

pequeña,

• Un trabajo así no permite al lector decidir por ellos mismos si ρ-value es suficientemente pequeño para rechazar o no aceptar H_0. Por ejemplo: Si un lector cree que H_0 no debe rechazarse a menos que $\rho < 0.01$ entonces informar que solamente que $\rho < 0.05$ no permite al lector determinar si rechaza o no acepta o acepta o no rechaza H_0,

• Valores pequeños de ρ-value señalan que H_0 es improbable que sea verdadera, tal que es tentador pensar que ρ-value representa la probabilidad que H_0 sea verdadera. No obstante, la verdad o falsedad de H_0 no se puede cambiar mediante la repetición del experimento; por lo tanto, no es correcto hablar de "probabilidad" que H_0 sea verdadera,

• La clase de probabilidad que analiza si no se rechaza o no se acepta H_0, útil solamente cuando se aplica a resultados que pueden resultar en formas diferentes cuando se repiten experimentos debido a que para definir el ρ-value como probabilidad de observar un valor extremo de un estadístico como $\bar{X}$, pues su valor podría ser diferente si el experimento se repite, se llama probabilidad frecuentista (la frecuencia de un suceso en una muestra se define como el cociente entre número de veces que ha ocurrido el suceso en la muestra y el tamaño de la misma.

• Empíricamente, se observa que al ir aumentando

el tamaño de una muestra, la frecuencia de sucesos tiende a estabilizarse de un número fijo tal que se ha denominado como ley de estabilidad de frecuencias o ley única del azar y ese número ideal, límite que alcanzaría la frecuencia de un suceso si se obtuviera una muestra infinita del experimento es el primer concepto de probabilidad de un suceso),

• La probabilidad subjetiva calcula la probabilidad que un enunciado, como , sea verdadero y es importante en la teoría de Estadística Bayesiana (es un subconjunto del campo de la estadística en la que la evidencia sobre el verdadero estado del mundo se expresa en términos de grados de creencia o, más específicamente, las probabilidades bayesianas. Es sólo una de una serie de interpretaciones de la probabilidad y hay otras técnicas estadísticas que no se basan en "grados de creencia".

La inferencia bayesiana es un enfoque de la inferencia estadística, que es distinta de la inferencia frecuentista. Se basa específicamente en el uso de probabilidades bayesianas al resumir las pruebas.

La formulación de modelos estadísticos para su uso en la estadística bayesiana tiene la característica adicional, no está presente en otros tipos de técnicas estadísticas, que requiere la formulación de un conjunto de distribuciones previas para los parámetros desconocidos, sus consideraciones habituales en diseño de experimentos se extienden en el caso de diseño

Bayesiano de experimentos para incluir la influencia de las creencias anteriores y requiere ciertas técnicas computacionales modernas para la inferencia bayesiana, como diversos tipos de técnicas Monte Carlo de cadenas de Markov) y

• Si un resultado tiene un pequeño se dice que es "estadísticamente significativo", representa "importante" y, en consecuencia, resulta tentador pensar que resultados estadísticamente significativos siempre son importantes. No obstante, a veces este tipo de resultados no tienen importancia científica o práctica.

Entonces, un resultado puede ser estadísticamente significativo sin ser lo suficientemente grande para tener importancia práctica, pues una diferencia es estadísticamente significativa cuando es grande comparada con su o e, incluso, cuando o es muy pequeña puede ser estadísticamente significativa.

Por lo tanto, no mide la significancia práctica, sino el grado de confianza que se puede tener que el valor verdadero es muy diferente del valor especificado por tal que cuando es pequeño se tiene la confianza que el valor verdadero es, en verdad, muy diferente, pero no implica que la diferencia sea lo bastante grande para que tenga importancia práctica".

• (Lind, Marchal, & Mason, 2006) sostienen "en una prueba de hipótesis sólo se puede tomar una de dos decisiones: aceptar o rechazar $H_{0.}$ En lugar de

"aceptar" H_0 algunos investigadores prefieren enunciar la decisión como "No se rechaza H_0", "No se puede rechazar H_0" o, también, "Los resultados muestrales no permiten rechazar H_0. En pruebas de hipótesis ρ-value, proporciona la probabilidad de obtener, suponiendo que H_0 sea verdadera, un valor muestral estadístico de prueba tan extremo, por lo menos o más extremo, como el obtenido tal que ρ-value compara la probabilidad con el nivel de significancia". Por lo tanto, si no se acepta H_0, el siguiente paso es efectuar la prueba de significancia entre medias de tratamientos con el fin de conocer cuál o cuáles son estadísticamente mejores.

- (Triola, 2004) afirma "estadístico de prueba es un valor calculado a partir de datos muestrales, que se usa para tomar la decisión sobre rechazo de H_0. Por lo tanto, sirve para determinar si existe evidencia significativa en contra de H_0. ρ es la probabilidad de obtener un valor del estadístico de prueba que sea al menos tan extremo como el que representa a datos muestrales, suponiendo que H_0 es verdadera.

Algunos libros dicen "aceptar H_0" en vez de "no rechazar H_0", pero no se está probando H_0, sea que use el término "aceptar" o "no rechazar", tal que únicamente se afirma que la evidencia muestral no es lo suficientemente fuerte como para justificar el rechazo de H0. El término "aceptar" es un poco confuso, pues parece indicar incorrectamente que H_0 ha sido aprobada tal que la frase "no rechazar" indica correctamente que

la evidencia disponible no es lo suficientemente fuerte para justificar el rechazo de H_0".

Usualmente, en ejercicios se pide demostrar que $t^2=F$ tal que $t = \frac{\hat{\beta}_2}{ee(\hat{\beta}_2)}$ quemuestra que las pruebas Student y Fisher son equivalentes.

Observación.

Si $\beta_2=b\neq 0$, F sigue una ley de Fisher descentrada con igual grados de libertad y parámetro de descentramiento $\delta_2=(b^2_2)/(Var(\beta^{\wedge}_2)$. La potencia de prueba es:

$$\eta(b) = Pr(F \leq F_{(1,n-2),}\alpha|\beta_2 = b)$$

4.3 Coeficiente de determinación

Es otra medida de relación entre las variables llamada coeficiente de determinación R^2. Se debe a que da mayor fuerza de interpretación a la relación entre las variables. Además, es un buen indicador de "calidad" de regresión, pero no es determinante ni suficiente para decidir sobre la adecuación del modelo. Su uso es muy difundido, pero en general inapropiado pues no tienen presentes las limitaciones y alcances de este indicador. Además, no estudian residuos.

4.3.1 Coeficiente

Una regresión será "buena" si la variabilidad explicada por regresión es relativamente alta respecto a variabilidad total de Y; es decir, si SEC≈STC. La proporción de variabilidad explicada por regresión se

mide con coeficiente de determinación.

Definición. Se llama coeficiente de determinación R^2 al cociente:

$$R^2 = \frac{SEC}{STC}$$

Teorema. Si la regresión tiene término constante o si $\bar{y}=0, 0 \leq R^2 \leq 1$.

Demostración.

Se sabe que si la regresión tiene término constante o si $\bar{y}=0$, entonces:

$$STC=SEC+SRC$$

Tal que, la suma de cuadrados es positiva y el cociente de dos números positivos es positivo. Si la regresión no incluye el término constante i $\bar{y} \neq 0$ no existe razón para que R^2 esté entre 0 y 1.

Interpretación. Se multiplica por 100 y se interpreta en términos de porcentaje tal que es el porciento de variabilidad explicada por la regresión con respecto a variabilidad total. La definición e interpretación de R^2 tiene sentido únicamente cuando la regresión lineal incluye el término constante o los datos han sido centrados y, por consecuencia, no hay término constante en la regresión.

Es importante que R^2 sea cercano a 1, pero no es determinante para calidad de regresión. Su calidad

se decide en función de todos sus indicadores, como coeficiente de determinación, razón F, signos esperados, rangos esperados para estimadores de parámetros β_j, pruebas de hipótesis sobre parámetros, gráficos de residuos y pruebas de hipótesis sobre residuos. Se puede demostrar que $F=(n-2)(R^2/(1-R^2))$ tal que R^2 será significativo si F lo es. El coeficiente puede ser "bajo" y la regresión significativa.

El coeficiente R^2 se puede usar en comparaciones de dos modelos de regresión tal que se prefiere el modelo con mayor valor de R^2. Para comparar dos modelos con el coeficiente R^2 es importante que la variable dependiente sea la misma.

Por ejemplo $y_i=\beta_0+\beta_1 x_i+u_i$ i $Ln(y_i)=\beta_0+\beta_1 x_i+u_i$ no son comparables con R_2. En primer caso R2 es propoción de variabilidad de Y explicada por X mientras que en segundo caso es la variabilidad de Z=Ln(Y) explicada por la misma variable. Este coeficiente puede ser fuertemente afectado por datos atípicos.

Coeficiente de Determinación r^2 (dos variables) ó R^2 (regresión múltiple) como medida de "bondad de ajuste". La bondad del ajuste de la línea de regresión de un conjunto de datos refiere a cuán "bien" se ajusta la línea de regresión a los datos. Con base en la siguiente figura, es claro que si todas las observaciones caen en la línea de regresión se obtiene un ajuste "perfecto", pero rara ocasión se presenta:

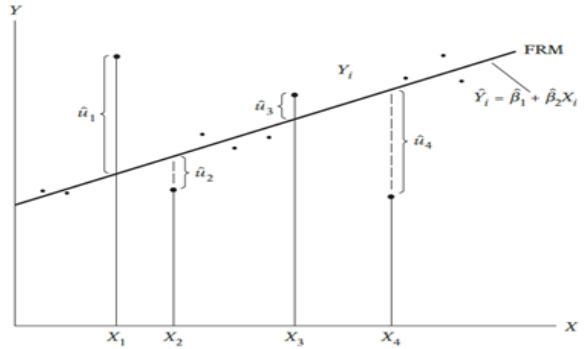


Figura 14. Bondad del ajuste de la línea de regresión de un conjunto de datos refiere a cuán "bien" se ajusta la línea de regresión a los datos

En general, hay algunas $\hat{u}_i$ positivas y algunas $\hat{u}_i$ negativas. Se tiene la esperanza que estos residuos alrededor de la línea de regresión sean lo más pequeños posibles. Entonces, el coeficiente de determinación r^2 (dos variables) ó R^2 (regresión múltiple) es una medida comprendida que indica cuán bien se ajusta la línea de regresión muestral a los datos. Una explicación del significado de heurística de r^2 ó R^2 en términos gráficos, conocida como Diagrama de Venn, Euler o Ballentine (r^2:a)=0;f) r^2=1) es:

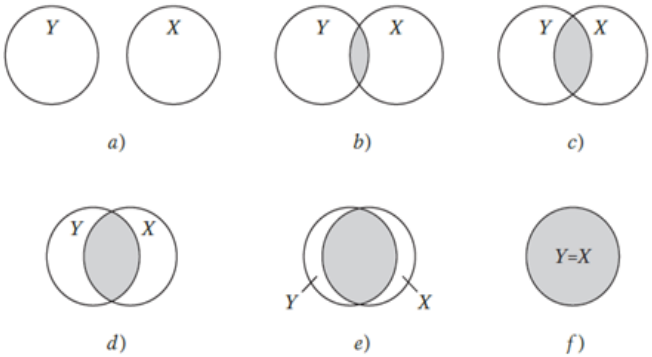


Figura 15. Diagrama de Venn, Euler o Ballentine

El círculo Y representa la variación en variable explicada, dependiente o endógena Y, el círculo X es la variación en la variable explicativa, independiente o exógena X (Varianza$_{\text{(Suma de cuadrados dividida por grados de libertad apropiados)}}$=Variación$_{\text{(Suma de cuadrados de desviaciones de una variable respecto a su media)}}$/Gl), mientras que la intersección de círculos (área sombreada) señala la media en que la variación en Y se explica por variación en X, como regresión de MCO.

A mayor medida de intersección, mayor será la variación en Y que se explica por X, pues a medida que va de izquierda a derecha, el área de intersección aumente o, en otras palabras, hay una proporción cada vez mayor de la variación en Y explicada por X.

Entonces, r^2 ó R^2 es una medida numérica de esta intersección. Por lo tanto, cuando no hay intersección r^2 ó R^2=0, cuando es completa r^2 ó R^2=1, entendida como 100% de variación de Y se explica por X y, por lógica, r^2 ó

R^2 varía ±1. El cálculo de r^2 ó R^2 se hace de la siguiente forma, pues si:

$Y_i=Y_{\hat{}i}+u\hat{}_i$ o, en términos de desviación, $y_i=y_{\hat{}i}+u\hat{}_i$

Donde se emplean ecuaciones $y_i=\beta\hat{}_2 x_i+u\hat{}_i$ y $y\hat{}_i=\beta\hat{}_2 x_i$ tal que al elevar al cuadrado $y_i=y\hat{}_i+u\hat{}_i$ en ambos lados y sumar sobre la muestra se obtiene:

$$\sum y_i^2=\sum \hat{y}_i^2+\sum \hat{u}_i^2+2\sum \hat{y}_i\hat{u}_i=\sum \hat{y}_i^2+\sum \hat{u}_i^2=\hat{\beta}_2^2\sum x_i^2+\sum \hat{u}_i^2$$

Pues $\sum y\hat{}_i\hat{u}_i=0$ y $y\hat{}_i=\beta\hat{}_2 x_i$. Las diversas sumas de cuadrados en ecuación anterior se describen de la forma siguiente:

$$\sum y_i^2=\sum (Y_i-\bar{Y})^2$$

Hace referencia a variación total de valores reales de Y respecto de su media muestral, denominada Suma de Cuadrados Total (SCT).

$$\sum \hat{y}_i^2=\sum (\hat{Y}_i-\bar{Y})^2=\sum (\hat{Y}_i-\bar{Y})^2=\hat{\beta}_2^2\sum x_i^2$$

Refiere a la variación de valores de Y estimador alrededor de su media $(Y_{\hat{}i}=\bar{Y})$, que apropiadamente puede llamarse Suma de Cuadrados por Regresión, debida a variables explicativas, independientes o exógenas, o, simplemente, Suma de Cuadrados explicada (SCE). Adicionalmente, $\sum u_i^2$ es variación residual o variación no explicada de valores de Y alrededor de la línea de regresión o, también, llamada Suma de Cuadrados de Residuos (SCR). Así, ecuación:

$$\sum y_i^2=\sum \hat{y}_i^2+\sum \hat{u}_i^2+2\sum \hat{y}_i\hat{u}_i=\sum \hat{y}_i^2+\sum \hat{u}_i^2=\hat{\beta}_2^2\sum x_i^2+\sum \hat{u}_i^2\Rightarrow$$

$$\therefore SCT=SCE+SCR$$

Esta ecuación muestra que la variación total en valores observados alrededor del valor de su media puede dividirse en dos partes, una atribuible a la línea de regresión y otra a fuerzas aleatorias, pues no todas las observaciones caen sobre línea ajustada. Geométricamente es:

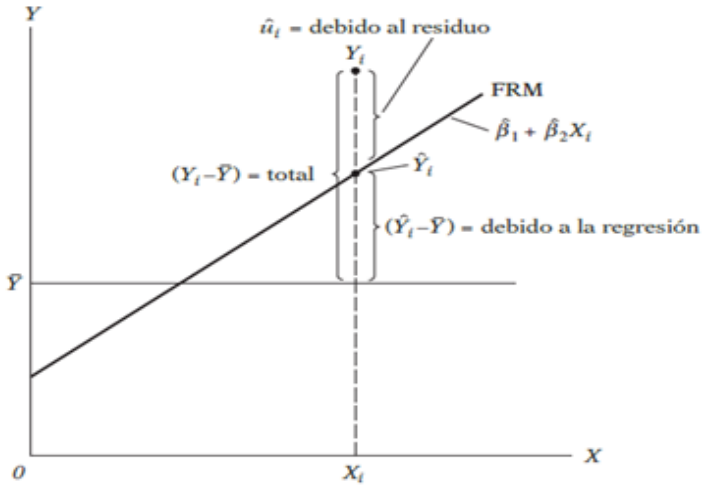


Figura 16. Variación total en valores observados alrededor del valor de su media puede dividirse en dos partes, una atribuible a la línea de regresión y otra a fuerzas aleatorias

Al dividir la ecuación SCT=SCE+SCR entre SCT en ambos lados, se obtiene:

$$\frac{\cancel{SCT}}{\cancel{SCT}} = \frac{SCE}{SCT} + \frac{SCR}{SCT} \Rightarrow 1 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} + \frac{\sum \hat{u}_i^2}{\sum(Y_i - \bar{Y})^2}$$

Ahora bien, se define r^2:

$$r^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{SCE_{(Error)}}{SCT_{(Total)}}$$

ó

$$r^2 = 1 - \frac{\sum \hat{u}_i^2}{\sum(Y_i - \bar{Y})^2} = 1 - \frac{SCR_{(Regresión)}}{SCT_{(Total)}}$$

La cantidad definida de esta manera se conoce como "Coeficiente de Determinación Muestral", entendida como la medida más común de bondad de ajuste de una lineal de regresión o, verbalmente, r^2 mide la proporción o por ciento de variación total en Y explicada por el modelo de regresión. Asimismo, r^2 observa dos propiedades:

1. Es una cantidad no negativa, pues al elevar al cuadrado cualquier número, sea positivo o negativo, el resultado es positivo (ley de signos).

2. Sus límites son $0 \leq r^2 \leq 1$. $r^2=1$ indica un ajuste perfecto; es decir, $\hat{Y}_i = Y_i$ por cada i. Por ortro lado, $r^2=0$ significa que no hay una relación alguna entre la variable regresada, dependiente, explicada o endógena y variable regresora, independiente, explicativa o exógena (coeficiente de pendiente o de variable exógena $\hat{\beta}\_i=0$). En este caso, según $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i = (\bar{Y} - \hat{\beta}_2 \bar{X}) + \hat{\beta}_2 X_i = Y$

$\bar{} + \hat{\beta}_2 (X_i - \bar{X})$, $\hat{Y}_i = \hat{\beta}_1 = \bar{Y}$; es decir, la mejor predicción de cualquier valor de Y es simplemente el valor de su media. En consecuencia, la línea de regresión será horizontal al eje X.

No obstante, r^2 puede calcularse directamente a partir de la definición o, más rápido, con la fórmula siguiente:

$$r^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = 1 - \frac{\sum \hat{u}_i^2}{\sum(Y_i - \bar{Y})^2} = 1 - \frac{SCR_{(Regresión)}}{SCT_{(Total)}} = \frac{SCE_{(Error)}}{SCT_{(Total)}} = \frac{\sum \hat{y}_i^2}{\sum y_i^2} = \frac{\hat{\beta}_2^2 \sum x_i^2}{\sum y_i^2}$$
$$= \hat{\beta}_2^2 \left(\frac{\sum x_i^2}{\sum y_i^2}\right)$$

Si se dividen el numerado y denominador de la ecuación anterior por el tamaño n muestral (n-1 si la muestra es pequeña), sabiendo que S_x^2-S_y^2 son varianzas muestrales X-Y respectivamente, se obtiene:

$$r^2 = \hat{\beta}_2^2 \left(\frac{\frac{\sum x_i^2}{n}}{\frac{\sum y_i^2}{n}}\right) = \hat{\beta}_2^2 \left(\frac{S_x^2}{S_y^2}\right)$$

Como $\hat{\beta}_2 = (\sum X_i Y_i)/(\sum x_i^2)$, la ecuación $\hat{\beta}_2^2 (\sum x_i^2)/(\sum y_i^2)$ también se expresa:

$$r^2 = \frac{(\sum X_i Y_i)^2}{\sum x_i^2 \sum y_i^2} = \frac{(\sum y_i \hat{y}_i)^2}{(\sum y_i^2)(\sum \hat{y}_i^2)}$$

Con la definición de r^2, $SCE_{(Error)}$ y $SCR_{(Regresión)}$ explicadas antes, se expresan así:

$$SCE_{(Error)} = r^2 * SCT_{(Total)} = (r^2)(\sum yi^2)$$

Por lo tanto:

$$SCR_{(Regresión)} = SCT_{(Total)} - SCE_{(Error)} = SCT_{(Total)} \left(1 - \frac{SCR_{(Regresión)}}{SCT_{(Total)}}\right) = \left(\sum y_i^2\right)(1 - r^2)$$

Por lo tanto, se escribe:

$$SCT_{(Total)} = SCE_{(Error)} + SCR_{(Regresión)}$$

ó

$$\sum y_i^2 = r^2 \sum y_i^2 + (1 - r^2)\left(\sum y_i^2\right)$$

4.3.2 Coeficiente de correlación

Si se tienen dos variables aleatorias, una medida de la relación que existe entre ellas es el coeficiente de correlación ρ. Para determinar si existe una relación lineal entre las variables productora y de respuesta se utiliza el coeficiente de correlación lineal de Pearson, denotado por r, que se define por:

$$r = \frac{SC_{xy}}{\sqrt{SC_{xx}\, SC_{yy}}} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sqrt{(\sum_{i=1}^{n} x_i^2 - n\,\bar{x}^2)\,(\sum_{i=1}^{n} y_i^2 - n\,\bar{y}^2)}}$$

Incluso, dada una pareja de variables aleatorias (X,Y) se define el coeficiente de correlación:

$$\rho = \frac{Cov(X,Y)}{\sqrt{[Var\,(X) Var\,(Y)]}} = \frac{Cov(X,Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

Tal que Cov(X,Y)=E[(X-E(X)(Y-E(Y))].

Teorema. Para ∀(X,Y) par de variables aleatorias, $|\rho| \leq 1$.

Teorema. Si las variables X,Y son independientes tal que ρ=0. Sin embargo, el recíproco de este teorema es, en general, falso. Si la distribución conjunta es normal, entonces independencia es equivalente a no correlación.

Se puede demostrar que si las parejas $(X_i,Y_i)_{i=1,2,3,4,\ldots,n)}$ son independientes, igualmente distribuidas con ley normal tal que el coeficiente de correlación (empírico) definido:

$$r = \frac{\sum_i (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_i (y_i - \bar{y})^2 \sum_i (x_i - \bar{x})^2}}$$

Es estimador de máxima verosimilitud de ρ. Además, si:

$$z = \tanh^{-1}(r) = \frac{1}{2} Ln\left(\frac{1+r}{1-r}\right)$$

Se tiene:

$$\sqrt{n-3}(z - \tanh^{-1}(\rho)) \rightarrow_{Ley} N(0,1)$$

Esto indica que para suficientemente grande:

$$z \approx N\left(\tanh^{-1}(\rho), \frac{1}{n-3}\right)$$

Se puede usar este resultado asintótico para estimar intervalos de confianza para ρ o para contrastar hipótesis. Sin embargo, el coeficiente de correlación r (√ r^2) es la cantidad estrechamente relacionada con el concepto de Coeficiente de Determinación r^2 (dos variables) ó R^2 (regresión múltiple) como medida de "bondad de ajuste".

El **Coeficiente de Correlación Muestral** explica el grado de asociación, mide la fuerza o grado de asociación lineal, entre dos variables. Su cálculo es a partir de:

$$r = \pm\sqrt{r^2} = \frac{\sum X_i Y_i}{\sqrt{(\sum x_i^2)(\sum y_i^2)}} = \frac{n\sum X_i Y_i - [(\sum X_i)(\sum Y_i)]}{\sqrt{[n\sum x_i^2 - (\sum X_i)^2][n\sum y_i^2 - (\sum Y_i)^2]}}$$

El Coeficiente de Correlación Poblacional ρ (letra griega minúscula Rho o ro) es:

$$\rho = \frac{\text{Cov}(X,Y)}{\sqrt{[\text{Var}(X)\text{Var}(Y)]}} = \frac{\text{Cov}(X,Y)}{\sqrt{\sigma_X^2\sigma_Y^2}} = \frac{\text{Cov}(X,Y)}{\sigma_X\sigma_Y}$$

Entonces, ρ es una medida de Asociación Lineal entre dos variables y su valor se sitúa entre ±1, donde -1 indica una perfecta asociación negativa y +1 es una perfecta asociación positiva. Con base en lo anterior, se deduce:

$$\text{Cov}(X,Y) = \rho(\sigma_X\sigma_Y)$$

Algunas propiedades de son:

• Tener signo positivo o negativo. según el signo del término en numerador de $r = \pm\sqrt{r^2} = \frac{\sum X_i Y_i}{\sqrt{(\sum x_i^2)(\sum y_i^2)}} = \frac{n\sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{[n\sum x_i^2 - (\sum X_i)^2][n\sum y_i^2 - (\sum Y_i)^2]}}$ que mide la de dos variables.

• Se ubica entre límites -1 y +1; es decir, $-1 \le r \le 1$.

• Es simétrico por naturaleza; es decir, el Coeficiente de Correlación entre X y Y, denotado por r_{XY}.

• Es independiente del origen y escala; es decir, si se define $X_i^* = aX_i + C$ y $Y_i^* = bY_i + d$ tal que a > 0, b > 0, c y d son constantes. Entonces, r entre X* y Y* es igual a r entre variables originales X y Y.

• Si X y Y son estadísticamente independientes, pues dos variables aleatorias X y Y son estadísticamente independientes si y sólo si (⇔) f(x,y)=f(x) f(y); es decir, si Función de Densidad de Probabilidad conjunta se

expresa como el producto de las FDP marginales, el Coeficiente de Correlación ente ellas es cero, aunque no significa que dos variables sean independientes. En otras palabras, una correlación igual a cero no necesariamente implica independencia (siguiente figura h).

Sea X una variable discreta que toma valores diferentes $x_1, x_2, x_3, x_4, \ldots, x\_n$ tal que la función $f(x)=P(X=x_i)$ para $i=1,2,3,4,\ldots,n$ y $f(x)=P(X=x_i)=0$ para $x \neq x_i$ se denomina Función de Densidad de Probabilidad Discreta (FDP) de X, donde $P(X=x_i)$ indica probabilidad que la variable discreta X tome valor de x_i.

En cambio, Función de Densidad de Probabilidad de Variable Aleatoria Continua (FDP) afirma que X es una variable aleatoria continua, se dice que f(x) es FDP de X si satisfacen las condiciones que $f(x) \ge 0$, $\int_{-\infty}^{\infty} f(x)\,\delta x = 1$ y $\int_a^b f(x)\,\delta x = P(a \le x \le b)$ donde $f(x)\,\delta x$ es elemento probabilístico (probabilidad asociada a un pequeño interval de una variable continua) y $P(a \le x \le b)$ es probabilidad que X se encuentre en intervalo a a b tal que para una variable aleatoria continua, en contraste con una variable aleatoria discreta, la probabilidad que X tome valor específico es cero pues $\int_a^a f(x)\,\delta x = 0$, la probabilidad para tal variable solo se mide sobre un rango o intervalo dado, como (a,b):

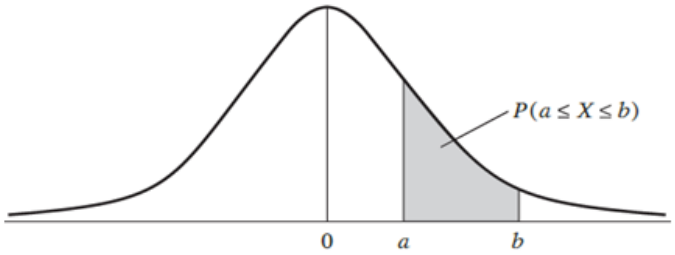


Figura 17. es probabilidad que se encuentre en intervalo a tal que para una variable aleatoria continua

La Función de Densidad de Probabilidad Conjunta Discreta indica que sean X y Y dos variables discretas tal que la función f(x,y)=P(X=x y Y=y)=0 cuando X≠x y Y≠y y da probabilidad (conjunta) que X tome valor de x y Y tome valor y.

La Función de Densidad de Probabilidad Individual o Marginal indica la relación con f(x,y),f(x) y f(y) se denominan funciones de densidad de probabilidad individuales o marginales, obtenidas mediante f(x)=$\sum_{}$ yf(x,y) (FDP marginal de X) y f(y)=$\sum_{x}$f(x,y) (FDP marginal de Y).

Finalmente, Función de Probabilidad Condicional del comportamiento de una variable condicional respecto a valores de otra (s) variable (s): f(x | y)=P(X=x | Y=y) tal que FPD condicional. Similarmente, f(y | x)=P(Y=y | X=x), FDP condicional de Y.

Las FDP condicionales se obtienen por f(x | y)=f(x,y)/ f(y) (FDP condicional de X) y f(y | x)=f(x,y)/(f(x)(FDP condicional de Y) tal que la FDP de una variable se expresa como la razón de FDP conjunta respecto de FDP marginal de otra variable (condicionante).

• Es una medida de Asociación Lineal o Dependencia Lineal solamente, su uso es la descripción de relaciones no lineales no tiene significado. Así, en la siguiente figura h, $Y=X^2$ es una relación exacta y r=0, pues es una expresión cuadrática (parábola positiva) y, en consecuencia, una relación no lineal.

• Es una medida de asociación lineal entre dos variables y no implica obligatoriamente una relación causa-efecto: "Una relación estadística por más fuerte y sugerente que sea nunca podrá establecer una conexión causal, por lo que ideas de causalidad provendrán de estadísticas externas y, finalmente, de una u otra teoría". Por lo tanto, "una relación estadística por sí misma no puede, lógicamente, implicar causalidad".

En el contexto de regresión, es una medida con más significado que , pues la primera indica la proporción de variación en variable dependiente, explicada, predicha, regresada, respuesta, endógena, resultado o controlada explicada por la (s) variable (s) independiente (s), explicativa (s), predictora (s), regresora (s), estímulo (s), exógena (s), covariante (s) o de control (s). En consecuencia, constituye una medida global del grado en

que la variación en una variable determina la variación de otra no tiene este valor y, además, la interpretación de en un modelo de regresión múltiple es de valor dudoso.

BIBLIOGRAFÍA

Agresti, A., & Finlay, B. (2018). *Statistical methods for the social sciences*. Pearson.

Bodie, Z., Kane, A., & Marcus, A. J. (2014). *Investments*. McGraw-Hill Education.

Cohen, L., Manion, L., & Morrison, K. (2013). *Research Methods in Education*. Routledge.

De Finetti, B. (1974). *Theory of Probability: A Critical Introductory Treatment*. John Wiley & Sons.

De Moivre, A. (1733). *The Doctrine of Chances:* A Method of Calculating the Probabilities of Events in Play. London, UK.

DeGroot, M. H., & Schervish, M. J. (2012). *Probability and Statistics*. Pearson Education.

DeVellis, R. F. (2012). *Scale development: Theory and applications* (Vol. 26). Sage Publications.

Devore, J. L. (2015). *Probability and Statistics for Engineering and the Sciences*. Cengage Learning.

Freund, J. E., & Walpole, R. E. (2019). *Mathematical Statistics with Applications*. Pearson.

García-Pérez, M. A. (2019). *Receiver Operating Characteristic* (ROC) Curves and Analysis for Binomial Detection (Vol. 3). CRC Press.

Hájek, A., & Hájek, P. (2011). *Interpretations of Probability*. Stanford Encyclopedia of Philosophy.

Hogg, R. V., & Tanis, E. A. (2019). *Probability and Statistical Inference*. Pearson.

Hogg, R. V., McKean, J., & Craig, A. T. (2018). *Introduction to Mathematical Statistics*. Pearson.

Howell, D. C. (2013). *Statistical methods for*

psychology. Cengage Learning.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press.

Kolmogorov, A. (1933). *Foundations of the Theory of Probability.* Chelsea Publishing.

Mendenhall, W., & Sincich, T. L. (2016). *Statistics for Engineering and the Sciences*. CRC Press.

Montgomery, D. C., Runger, G. C., & Hubele, N. F. (2016). Engineering Statistics. John Wiley & Sons.

Navidi, W. (2006). *Estadística para Ingenieros.* Colonia Desarrollo Santa Fe, Delegación Álvaro Obregón. México, D. F.: McGraw-Hill/Interamericana Editores, S.A. DE C.V.

Olmo, J. y Frias, D. (2000). «Capitulo 6. Métodos de dependencia. Regresión Lineal». En Técnicas de análisis de datos en investigación de mercados, 247-80. Madrid: Ediciones Pirámide.

Pascal, B., & Fermat, P. (1654). *Lettres de Mr. Pascal*, Contenant un Nouveau Mémoire sur les Jeux de Hazard. Paris, France.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research*. Explanation and prediction (2nd ed.). New York: Halt, Rinehart and Winston.

Peña, D. (1987).: *Estadística, modelos y métodos.* 2. Modelos lineales y series temporales Alianza Universidad.

Ross, S. M. (2010). *Introduction to Probability Models*. Academic Press.

Ross, S. M. (2019). *Introduction to Probability Models*. Academic Press.

Sánchez Vizcaino, G. (2000). «Capitulo 10. Métodos de dependencia. Regresión Logística». En Técnicas de análisis de datos en investigación de mercados, 431-67. Madrid: Ediciones Pirámide.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley & Son.

Shoeder et al. (1982). *Understanding regression analysis*: an introductory guide. Bervely Hills: Sage.

Silver, E. A., Pyke, D. F., & Peterson, R. (2018). *Inventory Management and Production Planning and Scheduling*. John Wiley & Sons.

Stevens, S. S. (1946). *On the theory of scales of measurement*. Science, 103(2684), 677-680.

Triola, M. F. (2004). *Estadística.* Naucalpan de Juárez, estado de México. México: Pearson Educación de México S. A. de C. V.

V. Abraira, A. Pérez de Vargas. *Métodos Multivariantes en Bioestadística*. Ed. Centro de Estudios Ramón Areces. 1996.

Wackerly, D. D., Mendehall, W. I., & Scheaffer, R. L. (2010). *Estadística Matemática con Aplicaciones.* Col. Cruz Manca, Santa Fe. México, D.F.: Cengage Learning Editores, S. A. de C. V., una Compañía de Cengage Learning, Inc.

Wang, C., Yin, J., & Wei, L. J. (2018). *Statistical Methods in Diagnostic Medicine*. John Wiley & Sons.

Wonnacott, T. H. and Wonnacott, R. J. (1981). *Regression: a second course in statistics.* New York: Wiley.

Nery Elizabeth García Paredes
Magister en Docencia y Currículo para la Educación Superior de la Universidad Técnica de Ambato (UTA); Magíster en Matemáticas de la Universidad Nacional del Litoral (UNL); Doctora en Ciencias de la Educación, mención: Física de la Escuela Superior Politécnica de Chimborazo (ESPOCH); Licenciada en Ciencias de la Educación , especialidad: Física y Matemáticas dela Universidad Técnica de Ambato (UTA); Docente Formador de Formadores de la Secretaria Superior, Ciencia, Tecnología e Innovación (SENESCYT); Docente Investigador de la Facultad de Contabilidad y Auditoría de la Universidad Técnica de Ambato(UTA; En su desarrollo profesional y académico ha desarrollado un total de diez artículos de investigación científica en diversas revistas indexadas. Actualmente se desempeña como docente en la carrera de Biotecnología de la Universidad Técnica de Cotopaxi; y docente la carrera de Administración de Empresas y Negocios Internacionales de la Pontificia Universidad Católica del Ecuador sede Ambato.

Alexander Fernando Haro Sarango
Docente Investigador del Instituto Superior Tecnológico España (ISTE); Magister en Sistemas de Información con mención en Inteligencia de Negocios y Analítica de Datos Masivos de la Universidad Estatal de Milagro (UNEMI); Licenciado Financiero en Universidad Técnica de Ambato (UTA);Investigador científico inscrito y reconocido por la Secretaría de Educación Superior de Ciencia, Tecnología e Innovación (SENESCYT – Ecuador) con Registro N.º REG-INV-22-05405.Durante su desarrollo profesional y académico ha desarrollado un total de cincuenta artículos de investigación científica en diversas revistas indexadas. Actualmente se desempeña como Docente de Administración Financiera y como Coordinador la carrera Tecnológica Superior Universitaria en Administración de Empresas e Inteligencia de Negocios en ISTE.

Lizeth Fernanda Silva Godoy
Máster Universitario en Ingeniería Matemática y Computación (UNIR – España); Licenciada en Ciencias de Educación, profesora de Ciencias Exactas (UNACH); Tecnóloga en Ciencias de Seguridad mención Aérea y Terrestre (ITSA); Docente de la Universidad de las Fuerzas Armadas ESPE sede Latacunga. En la actualidad se desempeña como Docente de Ciencias Exactas en Cálculo vectorial y Álgebra Lineal en los departamentos de Energía y Mecánica, Eléctrica y Electrónica, Ciencias de la Computación (ESPE-L), Latacunga -Ecuador

Fredin Fernando Pozo Parra
Máster Universitario en Ingeniería Matemática y Computación (UNIR – España); Licenciado en Ciencias de la Educación Mención Físico Matemático; Docente de la Universidad de las Fuerzas Armadas ESPE sede Latacunga; Docente de la Universidad Técnica Estatal de Quevedo: UTEQ; En su desarrollo profesional y académico ha desarrollado un total de tres artículos de investigación científica en diversas revistas indexadas; En la actualidad se desempeña como Docente de Algebra lineal y Cálculo Diferencial e Integral, de la Facultad de Ciencias Agrarias y Forestales (UTEQ), Quevedo – Ecuador

Stalin Gabriel Salguero Gualpa
Licenciado en Pedagogía de las Matemáticas y la Física de la Universidad Central del Ecuador (UCE); Docente y asesor independiente; En la actualidad se desempeña como asesor académico, capacitador en la institución privada ESCAM.

CASA EDITORA DEL POLO

ISBN: 978-9942-621-59-7
9 789942 621597