Proyecto PAPIME PE306125

Universidad Nacional Autónoma de México
Facultad de Economía
Dirección General de Asuntos del Personal Académico

Recursos didácticos digitales (manual y videos)

para estadística, probabilidad e introducción a la econometría

Autores

Dr. Edmar Ariel Lezama Rodríguez
Dr. Michel Eduardo Betancourt Gómez
Mtra. Laura Lázaro Rodríguez
Dra. Alicia Hernández Alfaro
C. Jaime Rico León
Becario del proyecto

Objetivo didáctico

Que el estudiantado comprenda y aplique los conceptos fundamentales de estadística, probabilidad e introducción a la econometría —incluyendo funciones de distribución, técnicas de muestreo y el método de Mínimos Cuadrados Ordinarios— para analizar datos reales y construir modelos básicos, mediante el uso de recursos didácticos digitales y ejercicios prácticos basados en la ENOE, con el fin de fortalecer su aprendizaje teórico-práctico.

1 de diciembre
2025

Contenido

Introducción

El presente manual es resultado del trabajo colaborativo entre personal académico y estudiantes de la UNAM, esfuerzo posible gracias al proyecto PAPIME PE306125, titulado “Recursos didácticos digitales (manual y videos) para estadística, probabilidad e introducción a la econometría”.

El texto está organizado en cinco capítulos, acompañados de videos tutoriales; responde a la experiencia docente acumulada durante los últimos diez años en el sistema escolarizado de la Facultad de Economía —particularmente en los cursos de primero, cuarto y quinto semestre—, la cual hemos identificado la dificultad recurrente del estudiantado para diferenciar conceptos fundamentales de estadística y probabilidad.

Esta misma experiencia ha permitido constatar que una comprensión sólida de dichos conceptos, así como el adecuado dominio de las funciones de distribución de variables aleatorias discretas y continuas, tiene un impacto directo y positivo en el aprendizaje de los principios básicos de la econometría, especialmente en lo relativo a los Mínimos Cuadrados Ordinarios.

Es importante destacar que los temas abordados en este material didáctico se acompañan de una base de datos diseñada específicamente para este fin, así como de videos tutoriales que facilitan la aplicación práctica de los contenidos curriculares mediante ejercicios guiados.

En cuanto a la base de datos, su construcción se realizó a partir de la Encuesta Nacional de Ocupación y Empleo (ENOE), levantada de manera trimestral desde 2005 por el Instituto Nacional de Estadística y Geografía (INEGI). Dado que la ENOE contiene información sociodemográfica y laboral, permite no solo aplicar los conocimientos teóricos y prácticos adquiridos en los

cursos de estadística y probabilidad, sino también constituye un insumo idóneo para la elaboración de los primeros modelos de Mínimos Cuadrados Ordinarios con orientación hacia el análisis económico.

Respecto a los videos que acompañan el texto y los datos, consideramos que representan un complemento formativo de gran valor, al ofrecer explicaciones centradas en aspectos prácticos y en problemáticas contemporáneas, tales como la brecha de género en los salarios, la oferta de horas en el mercado laboral mexicano, la relación entre escolaridad y salario, y la dualidad entre informalidad y formalidad, entre otros ejemplos.

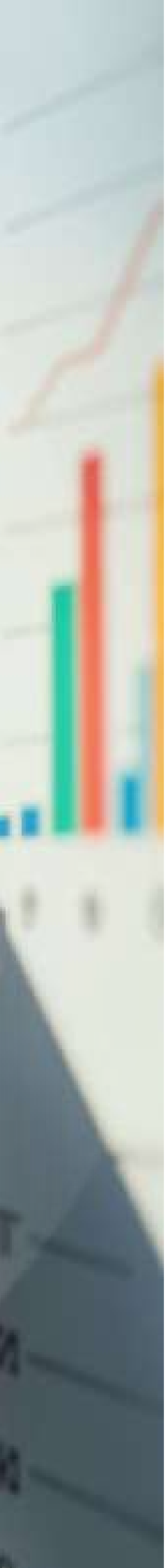
Confiamos en que este conjunto de materiales permitirá al estudiantado de la Licenciatura en Economía no sólo adquirir los conocimientos necesarios, sino aplicarlos de manera pertinente en diversos contextos sociales a lo largo de su formación profesional.

El video 1 que se muestra a continuación, además de dar la bienvenida al curso explica a manera de resumen los objetivos que persigue el presente manual y resaltar la importancia de dominar conceptos de estadística y probabilidad como introducción a econometría en cursos de economía a nivel licenciatura.

1

CAPÍTULO

Diferencia entre la estadística y la probabilidad: conceptos básicos



En este capítulo abordaremos las definiciones clásicas de estadística y probabilidad, así como las diferencias clave entre ellas.

El objetivo de comenzar definiendo y marcando las diferencias radica en el hecho de que muchos estudiantes de educación media superior y superior llegan a utilizar ambos términos como sinónimos, cuando en la práctica se refieren a conceptos distintos.

Resulta de suma importancia conocer la diferencia entre estas ramas de las matemáticas, ya que permite tener una mejor comprensión de la teoría económica y de la forma en que se pueden plantear distintos modelos económicos o de corte social.

1.1 Definición de ESTADÍSTICA

Al hacer una revisión bibliográfica sobre el concepto de estadística, seleccionamos algunas definiciones de diversos autores. Porras (2017) señala que es el conjunto de métodos para obtener, presentar y analizar observaciones numéricas, teniendo como objetivo final tomar decisiones o hacer generalizaciones de los fenómenos estudiados.

Rendón-Macías (2016) por su parte, define a la estadística como la disciplina que busca resumir datos de cualquier fenómeno en cuadros, tablas, figuras o gráficos con la finalidad de encontrar información puntual de distintos fenómenos y escenarios de la vida cotidiana.

Para Faraldo (2012), la estadística es un conjunto de técnicas numéricas y gráficas para describir y analizar un grupo de datos, sin extraer conclusiones (inferencias) sobre la población a la que pertenecen.

Otra definición de estadística es la que ofrece Coll Serrano (2023) al decir que se trata de una ciencia que se encarga de recoger, analizar, presentar e interpretar datos.

Para fines prácticos de este material se define a la estadística como la rama de las matemáticas que se encarga de recopilar información numérica a través de encuestas propias o realizadas por terceros (personas o instituciones) para ser analizada de forma simple o a través de medidas de tendencia central y no central; esta información puede ser agrupada en tablas o representada mediante gráficos para emitir una opinión fundamentada sobre un tema en particular (ver video 2).

https://youtu.be/gV5oaXQJJoY

1.2 Definición de PROBABILIDAD

Una vez que se ha definido a la estadística y su forma de trabajo, es necesario dar paso al concepto de probabilidad para así establecer las diferencias entre ambas ramas.

En lo que se refiere a probabilidad, Laplace (1814) ha hecho la construcción de una definición que resulta popular, al describirla como un cociente de los casos probables entre los casos posibles.

Los casos posibles son los resultados que se obtendrán en cada experimento. Por ejemplo, si consideramos el caso de un dado, se cuenta con seis posibles resultados (uno por cada una de las caras del dado); si se trata de una moneda, el resultado sería dos posibilidades.

En lo que se refiere a los casos probables, y siguiendo con el ejemplo de un dado, al tratar de obtener un número 3 después de hacer una tirada, la respuesta es una única ocasión, ya que sólo existe un número tres en un dado.

Dicho lo anterior y siguiendo la concepción de Laplace (1814), tenemos la siguiente fórmula:

$$Probabilidad = \frac{casos\ probables}{casos\ posibles}$$

Ejemplo: ¿Cuál es la probabilidad de obtener un número 3 al tirar un dado?

$$Probabilidad = \frac{1\ (solamente\ existe\ un\ número\ 3\ en\ el\ dado)}{6\ (existen\ 6\ caras\ en\ el\ dado\ y\ cada\ cara\ es\ un\ caso\ posible)}$$

$$Probabilidad = \frac{1}{6} \approx 0.166666667$$

Por lo tanto, la probabilidad de obtener un número 3 al tirar un dado es de aproximadamente 16.66%.

Al analizar la concepción que construyó Laplace (1814) y para fines prácticos de este trabajo, se define a la probabilidad como la razón o cociente de los eventos que se buscan probar entre los eventos totales (ver video 3).

1.3 DIFERENCIAS entre estadística y probabilidad

Una vez que hemos expuesto las definiciones de estadística y de probabilidad propuestas por distintas personas además de la propia, resulta importante señalar sus diferencias, lo cual ayudará a no utilizar como sinónimos ambos conceptos.

La diferencia sustancial entre la estadística y la probabilidad radica en que, mientras la primera busca organizar de manera sistemática información numérica, la segunda busca la posibilidad de ocurrencia de un evento en particular.

Si bien estadística y probabilidad no son sinónimos, la primera alimenta a la segunda; la información organizada de manera sistemática puede ser utilizada para encontrar la probabilidad de ocurrencia de algún evento en particular.

Pensemos en una encuesta sobre temas laborales realizada por el Instituto Nacional de Estadística y Geografía (INEGI; institución gubernamental en México encargada de recopilar, analizar y difundir información estadística y geográfica sobre el país). Dicha encuesta nos dirá cuántas personas están laborando y cuántas están desocupadas (información de carácter estadístico).

Pensemos en la población total como nuestro universo, y el conjunto de personas que están laborando y el conjunto de personas que están desempleadas o no laborando. Si quisiéramos conocer la probabilidad de encontrar a una persona desocupada, tendríamos que dividir al total de ese grupo

poblacional entre la cantidad total de personas ocupadas y desocupadas (evento probabilístico), es decir:

$$Probabilidad = \frac{Total\ de\ personas\ desocupadas}{Total\ de\ personas\ ocupadas\ y\ desocupadas}$$

Una vez que se conoce la definición de estadística y probabilidad, así como su diferencia, resulta importante mostrar la forma en cómo se representa la información obtenida y qué se puede hacer con ella.

En primer lugar, daremos un vistazo a los gráficos y a las medidas de tendencia central para explicar la parte de estadística; posteriormente, revisaremos la parte de conjuntos, lo cual nos ayudará a entender lo referente a probabilidad (ver video 4).

https://youtu.be/9F3mNj-McmY

1.4 NOCIONES BÁSICAS de estadística con sus fórmulas

La información estadística recabada puede ser representada de distintas formas, una de ellas es mediante gráficos y tablas (resumen visual), otra mediante su análisis a través de las medidas de tendencia central (medidas estadísticas) que pretenden resumir en un solo valor a un conjunto de datos y representan un centro en torno al cual se encuentra ubicado el conjunto de los datos, lo

cual nos lleva a tener a la media, la mediana y la moda como referencia de esas medidas de tendencia central (Quevedo, 2011).

Ahora que sabemos de la existencia de las medidas de tendencia central (MTC) es necesario definir a cada una de ellas.

1.4.1 Media aritmética

Una de las medidas de tendencia central más utilizadas en el campo de la estadística es la media aritmética; también se le conoce como promedio, media o valor esperado (ver video 5).

Esta medida usa la suma de los datos en un conjunto y los divide por el número total de datos de ese mismo conjunto. La fórmula se muestra a continuación:

$$\mathrm{Y} = \sum_{i=1}^{n} \frac{Yi}{n}$$

A manera de ejemplo, supongamos que en un curso de Economía tenemos 5 alumnos y las calificaciones finales registradas a cada uno son $5, 8, 10, 9, 10.$

Sustituyendo los valores en la ecuación anterior donde $Y_1 = 5;\ \ Y_2 = 8;\ Y_3 = 10;\ Y_4 = 9;\ Y_5 = 10,$ tenemos que:

$$Y_m = \frac{5 + 8 + 10 + 9 + 10}{5} = \frac{42}{5} \approx 8.4$$

Por lo tanto, la media aritmética de calificaciones del curso de Economía es ≈ 8.4

1.4.2 Moda

Dentro de las medidas de tendencia central, la moda hace referencia siempre al dato que más se repite; cuando en un conjunto de datos no existen valores que se repitan, decimos que no hay moda o el conjunto es amodal.

Si en el grupo de datos tenemos un único dato que es el más repetido, decimos que ese conjunto es unimodal, pero si tenemos dos datos o más que se repiten en la misma cantidad entre ellos, decimos que ese conjunto es multimodal.

A manera de ejemplo, usaremos la misma serie de notas que se presentaron la sección de media aritmética, es decir, las calificaciones de $5, 8, 10, 9, 10$ registradas en un curso de Economía.

La representación también puede hacerse como:

$$x_1 = 5;\ \ x_2 = 8;\ \ x_3 = 10;\ \ x_4 = 9;\ \ x_5 = 10$$

Al analizar los datos, se identifica que los registrados en x_3 y en x_5 son los que más se repiten y por lo tanto decimos que 10 es la moda en esta serie.

1.4.3 Mediana

La mediana se define como el valor intermedio entre un grupo de datos ordenados, es decir, es el valor que se encuentra a la mitad, lo cual hace que de un extremo de la serie queden los valores mayores de la serie y del otro, los menores.

Al calcular la mediana de nuestras calificaciones, acomodamos los datos de forma ascendente:

$$x_1 = 5; \quad x_2 = 8; \quad x_3 = 9; \quad x_4 = 10; \quad x_5 = 10$$

El valor de x_3 es el que se encuentra justo a la mitad, por lo que decimos que 9 es la mediana de esta serie de calificaciones.

Cabe aclarar que cuando el número de datos totales es un número par, por ejemplo 6, la mediana se obtiene del promedio de los dos datos centrales (promedio de los datos en las posiciones 4 y 5).

1.4.4 Media geométrica

Es una medida de tendencia central que se utiliza para mostrar cambios porcentuales en una serie de números positivos (ver video 5).

La fórmula correspondiente es:

$$Media\ Geométrica = \left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}}$$

A manera de ejemplo y continuando con la serie de calificaciones, decimos que la media geométrica es:

$$MG = \sqrt[5]{5 \times 8 \times 9 \times 10 \times 10}$$

$$MG = \sqrt[5]{36,000} \approx 8.109$$

$$MG \approx 8.11$$

En el siguiente video se aborda la diferencia entre los conceptos de media aritmética y media geométrica.

1.5 TEORÍA DE CONJUNTOS, las bases para entender a la probabilidad

Una vez que hemos definido los conceptos de estadística y probabilidad, así como sus diferencias, resulta importante introducir más conceptos y teorías que nos permitan entender para qué sirve la probabilidad.

Comenzaremos con lo referente a teoría de conjuntos, ya que esa es la manera en cómo representamos información para después transformarla en un modelo probabilístico o econométrico.

La teoría de conjuntos es una de las ramas más importantes en el área de matemáticas y tiene su origen en el trabajo realizado por Cantor (1874), el cual realizó otros desarrollos que permitieron consolidar tanto la definición como los preceptos de la teoría de conjuntos.

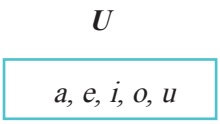
1.5.1 Definición de conjunto

En lo que se refiere a la definición, un conjunto es una colección de objetos o elementos y puede ser representado mediante un diagrama o dentro de corchetes o paréntesis; se presentan por letras mayúsculas y de forma particular se utiliza la letra U para denotar al universo de datos.

Si es entre corchetes, la forma de representar un conjunto universo (U) es:

$$U = \{ a, e, i, o, u \}$$

El conjunto anterior es una colección de letras que representa las vocales. Para el caso de un diagrama, la representación se muestra a continuación:



Por definición, el conjunto universo es el conjunto que contiene todos los elementos bajo estudio y se denota con la letra U o el símbolo Ω (omega). A partir de este conjunto pueden definirse subconjuntos que incluyen parte de dichos elementos.

Retomemos el mismo ejemplo previo donde el conjunto universo que contiene a las vocales podemos derivar un subconjunto A que contenga tres letras, tal como se muestra a continuación:

$$U = \{ a, e, i, o, u \}$$

$$A = \{ a, i, u \}$$

Por tanto, el conjunto A tiene únicamente tres vocales que fueron tomadas del conjunto universo.

A manera de nomenclatura matemática, generalmente todo conjunto o subconjunto se representa con una letra mayúscula iniciando con la A; los índices se denotan con minúsculas cursiva, iniciando con i, j, k, ... y los ejes en un plano cartesiano se denotan como x, y, z.

Una vez que hemos definido a un conjunto y la manera en que pueden ser representados, es importante mencionar que buena parte de los modelos estadísticos y econométricos son una representación de información a partir de una colección de datos delimitados por llaves, lo cual nos lleva a explicar la forma en cómo dicha información puede ser representada.

1.5.2 Notación de conjuntos

Un conjunto puede ser representado por extensión o comprensión; siguiendo con el ejemplo de las vocales.

Si es por extensión, la forma es: $U = \{ a, e, i, o, u \}$ donde se numeran todos y cada uno de sus elementos.

Si es por comprensión, la forma es: $U = \{ x \mid x \text{ es una vocal} \}$ se representan mediante expresiones matemáticas.

La representación de un conjunto ya sea por extensión o por comprensión, hace referencia al total de datos u objetos de estudio; vale la pena mencionar que resulta más práctico expresarlos por comprensión, sobre todo si nuestro conjunto tiene cientos o miles de objetos de estudio.

Por ejemplo, si mi intención es representar a toda la Población Económicamente Activa (PEA) de México en un conjunto por extensión, tendría que colocar el nombre de todas y cada una de las personas que componen a ese grupo.

Por otro lado, al expresar el conjunto por comprensión sólo hace falta referenciar a la PEA de México, delimitando así al conjunto tal como se muestra a continuación:

$$U = \{ x \mid x \text{ es la PEA de México} \}$$

$$A = \{ x \mid x \text{ es la Población Ocupada de México} \}$$

Vale la pena mencionar que el ejemplo presentado es una buena representación de cómo definir un conjunto por comprensión y un subconjunto que se deriva del conjunto universo.

El conjunto U contiene a toda la PEA, que por definición son todas las personas de 12 y más años que realizaron algún tipo de actividad económica (población ocupada), o que buscaron activamente hacerlo (población desocupada abierta) en los dos meses previos a la semana de levantamiento; por tanto, la PEA se segmenta en población ocupada y población desocupada abierta o desocupados activos (INEGI, 2025).

Si consideramos al conjunto A, el cual contiene a la Población Ocupada de México, sabemos que ese grupo poblacional pertenece a la PEA, razón por la cual, podemos afirmar que A es un subconjunto del conjunto U.

En lo que se refiere a nociones de la teoría de conjuntos, mostramos a continuación las definiciones básicas y sus formas de representación (ver videos 6 y 7).

- **Pertenencia**. Denotamos que el elemento x pertenece a A de la siguiente manera:

$$x \in A$$

- **No pertenencia**. Para decir que x no pertenece a A , o sea que el elemento x no es un elemento de la colección de A, lo denotamos de la siguiente manera:

$$x \notin A$$

- **Extensión**. Decimos que dos conjuntos son iguales si y solo si tienen los mismos elementos; lo denotamos de la siguiente manera:

$$x \in A \Leftrightarrow x \in B$$

El lema de extensión, nos indica que si un elemento y un conjunto de elementos pertenece a un conjunto A y ese mismo conjunto de elementos está contenido en B, entonces A y B son los mismos conjuntos.

- **Contención**. Sea A un conjunto contenido en B; cuando todos los elementos de A son también elementos de B, y se denota como:

$$A \subseteq B$$

- **Conjunto vacío.** Un conjunto vacío se define como aquel que no contiene ningún elemento y se denota de la siguiente manera:

$$A = \{ \emptyset \} = \{ \}$$

1.5.3 Operaciones entre conjuntos

- **Unión**. Sea A y B dos conjuntos, decimos que la unión de ambos conjuntos son el total de elementos que están en el conjunto A y los elementos del conjunto B.

$$A \cup B = \{x \mid x \in A \vee x \in B\}$$

El símbolo | denota una proposición matemática que debe ser cumplida por el elemento x, y se lee "tal que".

- **Intersección**. Sea A y B dos conjuntos, su intersección se define como:

$$A \cap B = \{x | x \in A \wedge x \in B\}$$

Sea x "tal que" x pertenezca al conjunto A y x pertenezca al conjunto B; es decir, la intersección de conjuntos se define como el conjunto de elementos que pertenecen al conjunto A y que también pertenecen al conjunto B.

- **Diferencia de conjuntos**. Sea A y B dos conjuntos; definimos la diferencia como:

$$A - B = \{x | x \in A \wedge x \notin B\}$$

- **Complemento**. Ahora imaginemos que A es algún subconjunto de U, por lo tanto:

$$A \subseteq U$$

Definimos la diferencia como $A - U$ y se denota como A^c

$$A\hat{}c = \{x \mid X \in U \wedge x \notin A\}$$

En la siguiente figura observamos como se vería el complemento entre conjuntos:

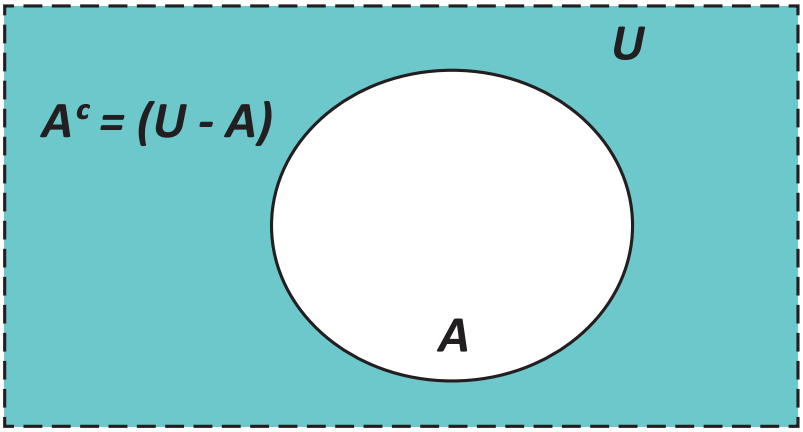


Figura 1.1 Reresentación del complemento. Fuente: elaboración propia.

Hasta aquí las definiciones básicas de conjuntos y su teoría, las cuales tienen una gran relevancia para la economía, en particular los conceptos de unión, intersección y complemento.

Pensemos en el siguiente ejemplo:

$$U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

$$A = \{1, 2, 6, 8, 10\}$$

$$B = \{1, 2, 6, 7, 8, 9\}$$

Donde U es un conjunto universo y los conjuntos A y B se derivan de dicho conjunto universal.

Si quisiera representar la unión entre el conjunto A y B, tendría lo siguiente:

$$A \cup B = \{1, 2, 6, 8, 10\} \cup \{1, 2, 6, 7, 8, 9\}$$

$$A \cup B = \{1, 2, 6, 7, 8, 9, 10\}$$

Como ya se mencionó, la unión de conjuntos hace referencia a los elementos que están en A y a los elementos que están en B; por esa razón en la unión vemos la totalidad de elementos que se encuentran en los conjuntos en cuestión, pero sin repetirse.

En el caso de la intersección, tenemos:

$$A \cap B = \{1, 2, 6, 8\}$$

En este caso, la intersección de conjuntos hace referencia a los elementos que están en A y también están en B, por esa razón en la intersección vemos los elementos que se repiten en ambos conjuntos.

Para poder representar el complemento del conjunto A con el conjunto U, tenemos lo siguiente:

$$U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

$$A = \{1, 2, 6, 8, 10\}$$

$$A^c = \{3, 4, 5, 7, 9\}$$

En este caso, el conjunto complemento hace referencia a aquellos elementos que faltan para llegar de nueva cuenta al conjunto universo.

Al conjunto A le falta el número 3, 4, 5, 7 y 9 para convertirse en el conjunto universo, por esa razón A^c tiene esos números.

Veamos la aplicación de conjuntos utilizando la base de datos de la ENOE generada para este manual (ver video 8).

1.6 EXPERIMENTOS determinísticos

En lo que se refiere a experimentos determinísticos, éstos se definen como aquellos que se realizan de una misma forma y con las mismas condiciones iniciales, y por lo tanto siempre se obtiene el mismo resultado.

Para el caso de los experimentos aleatorios, en ellos no se puede predecir el resultado final y debe cumplir con las siguientes condiciones:

- Se pueden repetir de manera indefinida.
- No se puede asegurar cuál será el resultado final, ya que esa es la esencia de un evento aleatorio.
- El resultado obtenido, pertenece a un conjunto de posibles resultados, el cual se conoce como espacio muestral.
- Cualquier subconjunto del espacio muestral es conocido como suceso aleatorio.

1.6.1 Evento seguro

En lo que se refiere al evento seguro, es aquel que siempre se verifica después de llevar a cabo el experimento aleatorio. Un ejemplo de evento seguro es aquel en el cual se obtiene una bola de color azul a partir de un recipiente que sólo tiene bolas azules.

1.6.2 Evento imposible

En lo que se refiere al evento imposible, es aquel que nunca se verifica como resultado del experimento aleatorio. Un ejemplo de evento imposible es obtener un número 8 al lanzar un dado que solo tiene seis caras.

Una vez presentadas las definiciones anteriores, estamos ya en condiciones de explicar qué es la probabilidad axiomática y qué usos tiene.

1.7 Probabilidad AXIOMÁTICA (Kolmogorov)

El concepto de probabilidad axiomática fue desarrollado por Kolmogorov (1933). Está definida por los siguientes axiomas:

1.7.1 Axiomas

- La probabilidad sólo puede tomar los valores comprendidos entre cero y uno:

$$0 \leq Probabilidad \leq 1$$

- La probabilidad del evento seguro es uno:

$$P(A) = 1$$

- La probabilidad de dos sucesos ajenos, disjuntos o mutuamente excluyentes es la suma de sus probabilidades respectivas.

Si

$$A \cap B = \emptyset$$

Entonces

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

Por lo tanto

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

A partir de los tres axiomas previos, es posible derivar las reglas que se consideran fundamentales en una función de probabilidad.

1.7.2 Reglas de función de probabilidad

- La probabilidad de la intersección de dos sucesos es menor o igual que la probabilidad de cada uno de los sucesos por separado, es decir:

$$P(A \cap B) \leq P(A) \quad y \quad P(A \cap B) \leq P(B)$$

- La probabilidad de unión de sucesos es mayor que la probabilidad de cada uno de los sucesos por separado:

$$P(A \cup B) \geq P(A) \quad y \quad P(A \cup B) \geq P(B)$$

- La probabilidad del suceso complementario del evento A es:

$$P(A^c) = 1 - P(A)$$

1.8 TEOREMA BÁSICO de probabilidad

A partir de los tres axiomas estudiados y de las reglas de función de probabilidad presentados, es posible desarrollar el teorema básico de probabilidad, el cual se muestra a continuación:

Si A y B son dos eventos cualesquiera en el espacio muestral S, entonces:

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

Si A, B y C son tres eventos cualesquiera de un espacio muestral S, entonces:

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

A manera de ejemplo, se plantea el siguiente ejercicio:

De un total de 100 personas, 40 estudian, 20 trabajan, y 10 estudian y trabajan. Si se elige a una persona al azar, calcular la probabilidad de que estudie o trabaje.

Sea el evento A: personas que únicamente estudian.

$$P(A) = \frac{40}{100}$$

Sea el evento B: personas que únicamente trabajan.

$$P(B) = \frac{20}{100}$$

Sea el evento ($A \cap B$): personas que estudian y trabajan.

$$P(A \cap B) = \frac{10}{100}$$

Por lo tanto:

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

$$P(A \cup B) \frac{40}{100} + \frac{20}{100} - \frac{10}{100} = \frac{50}{100}$$

Conclusión para este ejemplo:

Existe un 50% de probabilidad de encontrar a una persona que estudie o trabaje.

Una vez que se ha desarrollado la teoría de conjuntos y los conceptos básicos de probabilidad, podemos abordar con mayor certeza el tema de probabilidad condicional.

1.9 Probabilidad CONDICIONAL

La probabilidad condicional está definida como la probabilidad de que ocurra algún evento o suceso, dado que ya ocurrió algo previamente; es decir, el evento que se estudia está condicionado por un evento previo.

La probabilidad condicional se refiere a la relación entre el evento A y el evento B, siendo dos sucesos cualquiera donde la probabilidad de ocurrencia de B sea estrictamente mayor que cero. Matemáticamente se expresa como:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Lo cual se lee como la probabilidad de que ocurra el evento A, dado que ya ocurrió el evento B. Por esa razón es indispensable que la probabilidad de ocurrencia de B sea mayor que cero.

A manera de ejemplo se propone el siguiente ejercicio:

En una escuela de Economía, del total de estudiantes activos el 25% no aprobó matemáticas, el 15% no aprobó historia y el 10% reprobó historia y matemáticas.

Si se selecciona un estudiante al azar, calcular las siguientes probabilidades:

a. Si un estudiante reprobó historia, ¿cuál es la probabilidad de que haya reprobado matemáticas?

b. Si un estudiante reprobó matemáticas, ¿cuál es la probabilidad de que haya reprobado historia?

$$P(A) = Reprobar\ matemáticas$$

$$P(A) = \frac{25}{100}$$

$$P(B) = Reprobar\ historia$$

$$P(B) = \frac{15}{100}$$

$$Reprobar\ matemáticas\ e\ historia$$

$$P(A \cap B) = \frac{10}{100}$$

Para dar respuesta al inciso a), es necesario plantear que primero se debió reprobar historia, para posteriormente haber reprobado matemáticas, por tanto, la fórmula queda planteada de la siguiente manera:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{10}{100}}{\frac{15}{100}} = \frac{1{,}000}{1{,}500} \approx 0.666666667$$

Por tanto, la probabilidad de encontrar a un estudiante que haya reprobado matemáticas dado que reprobó historia previamente, es del 66.66%, aproximadamente.

Para el inciso b), el planteamiento es:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{10}{100}}{\frac{25}{100}} = \frac{1{,}000}{2{,}500} = 0.40$$

Por tanto, la probabilidad de encontrar a un estudiante que haya reprobado historia dado que reprobó matemáticas previamente, es del 40%.

Al realizar ambos ejercicios, vale la pena mencionar que el numerador de la fórmula siempre será la intersección de ambos eventos, mientras que el denominador hace referencia a la probabilidad que condiciona el evento.

Una vez que se ha explicado lo concerniente a la probabilidad condicional, es necesario pasar a la parte de eventos independientes.

1.10 EVENTOS independientes y dependientes

1.10.1 Eventos independientes

En el caso de los eventos independientes, estos se caracterizan por el hecho de que el resultado del segundo evento no se ve afectado ni influenciado por el primero; es decir la ocurrencia de un suceso no tiene impacto o influencia en la ocurrencia o no del otro.

Por tanto, los sucesos de A y B se dicen independientes si:

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

Lo que equivale a que:

$$P(A|B) = P(A)\ siempre\ y\ cuando\ P(B) > 0$$

O bien

$$P(B|A) = P(B)\ siempre\ y\ cuando\ P(A) > 0$$

Por tanto, si A y B son sucesos independientes, también lo son A y B^c.

Lo que demuestra que:

$$P(A \cap B^c) = P(A - A \cap B) = P(A) - P(A)P(B)$$

$$P(A \cap B^c) = P(A)[1 - P(B)] = P(A)P(B^c)$$

Como ejemplo, se muestra el siguiente ejercicio:

Si un dado se lanza dos veces, ¿cuál es la probabilidad de que en el primer lanzamiento salga el número 5 y en el segundo lanzamiento un número igual o menor que 3?

Considerando que el primer lanzamiento no influye en el segundo lanzamiento se dice que los eventos son independientes, por tanto:

$$Probabilidad\ de\ obtener\ el\ número\ 5 = P(A) = \frac{1}{6}$$

$$Probabilidad\ de\ tener\ un\ número\ igual\ o\ menor\ a\ 3 = P(B) = \frac{3}{6}$$

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

$$P(A \cap B) = \left(\frac{1}{6}\right)\left(\frac{3}{6}\right) = \frac{3}{36} \approx 0.083333333$$

Por tanto, la probabilidad de lanzar un dado y que en el primer lanzamiento salga el número 5 y en el segundo lanzamiento un número igual o menor que 3 es de aproximadamente 8.33%.

1.10.2 Eventos dependientes

En lo que se refiere a eventos dependientes, se define como aquellos donde la ocurrencia de un evento previo resulta determinante para que suceda el siguiente.

Como fórmula se tiene:

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

La anterior fórmula se lee como la probabilidad de encontrar dos eventos dependientes es igual al producto del primer evento (probabilidad de A) por la probabilidad condicional de esos dos eventos.

Como ejemplo, se desarrolla el siguiente ejercicio:

Una baraja tiene 52 cartas y dentro de ese total, existen 4 marcadas con el número 2. Si se toma una primer carta y no se devuelve a la baraja, para posteriormente tomar una segunda carta, ¿cuál es la probabilidad de que ambas cartas tengan el número 2?

Sin duda los eventos son dependientes, ya que el primero de ellos determina al segundo en el sentido de que pasamos de tener 52 a 51 cartas, además de que si en la primera carta que se toma se obtiene el número dos, afecta la probabilidad de volver a obtener el mismo valor en el siguiente turno, ya que inicialmente se podría tener cuatro veces al número 2 y ahora sólo se puede tener tres veces.

El resultado sería:

$$P(A \cap B) = P(A)P(A|B) = \left(\frac{4}{52}\right)\left(\frac{3}{51}\right) \approx 0.00452489$$

Por tanto, la probabilidad de sacar una carta con el número dos en un primer intento y repetir esa cifra en la segunda oportunidad, es de aproximadamente 0.45%.

1.11 TEOREMA de Bayes

Es nombrado así, en honor a Thomas Bayes, quien desarrolla la fórmula del teorema que lleva su nombre y se publica en 1763 en “Un ensayo para resolver un problema en la doctrina de las probabilidades”.

Actualmente, el Teorema de Bayes es una herramienta fundamental para resolver problemas en distintas ramas de la economía, sociología, ciencia política y administración pública, por solo citar algunas.

La definición es la de un proceso que nos sirve para calcular la probabilidad de un evento teniendo como referencia la ocurrencia de otros sucesos, los cuales están asociados a la probabilidad condicional.

La fórmula y demostración del Teorema de Bayes se muestra a continuación.

Partiendo de la probabilidad condicional tenemos que:

$$P(B_r|A) = \frac{P(B_r \cap A)}{P(A)}$$

Y del siguiente teorema:

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)(A|B_i)$$

Cuya demostración es:

$$A = (B_1 \cap A) \cup (B_2 \cap A) \cup \ldots \cup (B_k \cap A)$$

$$P(A) = P[(B_1 \cap A) \cup (B_2 \cap A) \cup \ldots \cup (B_k \cap A)]$$

Entonces:

$$P(A) = P(B_1 \cap A) + P(B_2 \cap A) + \ldots + P(B_k \cap A)$$

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)\, P(A|B_i)$$

Sustituyendo la última fórmula en correspondiente a la probabilidad condicional tenemos:

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B)P(A|B_i)}$$

Y como:

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

Por lo tanto:

$$P(B_r|A) = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B)P(A|B_i)}$$

Lo cual nos lleva a tener la fórmula del Teorema de Bayes.

A manera de ejemplo, se propone el siguiente ejercicio:

> Las máquinas A, B y C producen el 50%, 30% y 20% respectivamente del total de las mercancías de una fábrica manufacturera.
>
> Los porcentajes de producción defectuosa de estas máquinas son 3%, 4% y 5% respectivamente.
>
> Si se selecciona una mercancía al azar, ¿cuál es la probabilidad de que esa mercancía esté defectuosa y provenga de la máquina A?

Para resolver el problema, el primer paso es graficar un diagrama de árbol, como se muestra en la figura.

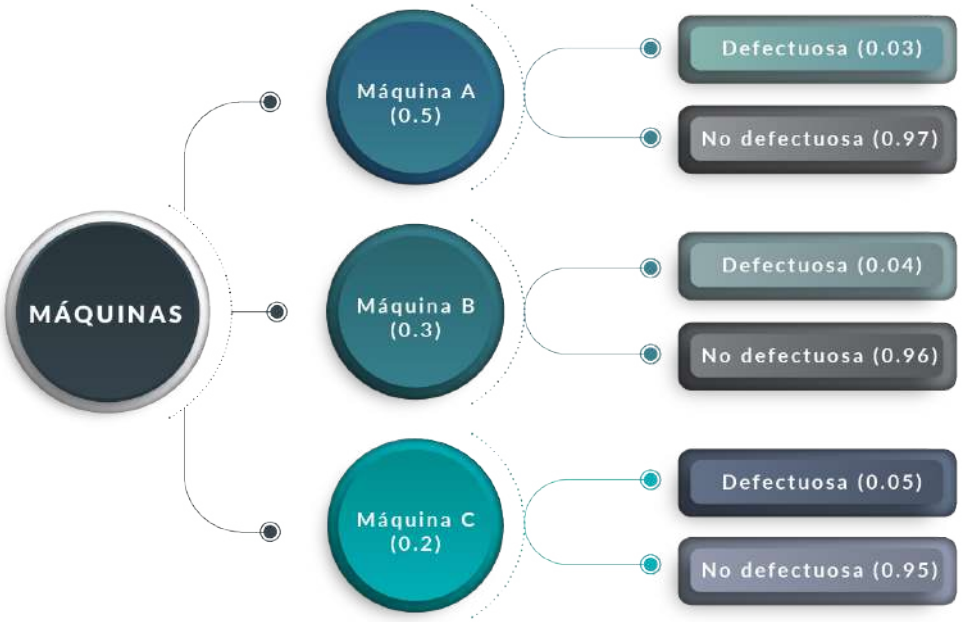


Figura 2.1
Diagrama de árbol de producción defectuosa.
Fuente: elaboración propia.

Este diagrama ejemplifica las tres máquinas del proceso productivo y qué porcentaje le corresponde a cada una respecto al total de la producción.

El último nivel del árbol muestra la probabilidad de encontrar una mercancía defectuosa y una no defectuosa para cada una de las máquinas, por lo que estamos ya en condiciones de sustituir los valores en la fórmula del Teorema de Bayes.

Por tanto, si necesitamos saber cuál es la probabilidad de elegir una mercancía que esté defectuosa y provenga de la máquina A, lo referente al numerador de la fórmula hará referencia a esa ruta, es decir, multiplicaremos la probabilidad de encontrar una mercancía de la máquina A (0.5) por la probabilidad de que esa máquina tenga mercancías defectuosas (0.03).

En lo que se refiere al denominador, es necesario multiplicar la probabilidad de producción de cada máquina por la probabilidad de que esa máquina produzca una mercancía defectuosa.

El resultado se muestra a continuación.

$$P(A|D) = \frac{(0.5)(0.03)}{(0.5)(0.03) + (0.3)(0.04) + (0.2)(0.05)}$$

$$P(A|D) = \frac{0.015}{0.015 + 0.012 + 0.01}$$

$$P(A|D) \approx 0.4054$$

Por tanto, la probabilidad de elegir al azar una mercancía defectuosa y que provenga de la máquina A es de aproximadamente 40.54% (ver video 9).

EJERCICIOS propuestos

1. Se desea conocer la ubicación de una fábrica de muebles, por lo que la probabilidad de que dicha fábrica se ubique en el municipio de A es del 70%, de que se ubique en el Municipio B es del 40% y de que se encuentre en ambos municipios es del 80%.¿Cuál es la probabilidad de que se localice en?

   a. Cualquiera de los dos municipios

   b. Ninguno de los municipios

2. En una compañía de 1500 empleados, 250 están en la sección de atención a clientes y 100 están en la sección de ventas al mayoreo. Calcular la probabilidad de que un empleado:

   a. No se encuentre en la sección de atención a clientes.

   b. No se encuentre en la sección de ventas al mayoreo.

   c. Se encuentre en la sección de atención al cliente o en la sección de ventas al mayoreo.

3. En una encuesta selectiva entre 180 personas que habían comprado un auto, 27 compraron un auto a gasolina y 72 uno eléctrico. Calcular la probabilidad de elegir a una persona que haya comprado un auto eléctrico, uno híbrido y ninguna de las opciones mencionadas.

4. La probabilidad de que un vuelo de programación regular despegue a tiempo es de 83%; la de que llegue a tiempo es 82% y la de que despegue y llegue a tiempo es del 78%. Calcular la probabilidad de que un vuelo elegido al azar:

   a. Llegue a tiempo dado que despegó a tiempo.

   b. Despegue a tiempo dado que llegó a tiempo.

5. De un total de 200 adultos, se tiene la siguiente información:

Tabla 1.1
Datos de escolaridad desglosada por sexo.
Fuente: elaboración propia.

Escolaridad	Hombre	Mujer
Primaria	38	45
Secundaria	28	50
Bachillerato	22	17

Si se selecciona aleatoriamente a una persona de ese grupo, calcular la probabilidad de que:

a. Sea hombre, dado que tiene escolaridad de nivel secundaria.

b. Sea mujer, dado que tiene escolaridad de nivel bachillerato.

6. La probabilidad de que un médico diagnostique en forma correcta una enfermedad es del 70%. Cuando el médico realiza un diagnóstico incorrecto la probabilidad de que un paciente presente una demanda es del 90%, ¿cuál es la probabilidad de que el médico haga un diagnóstico incorrecto y el paciente presente una demanda?

7. De la población estudiantil de cierta facultad, el 4% de los hombres y el 1% de las mujeres miden más de 180 cm. Además, el 60% de los estudiantes son mujeres. Ahora bien, si se selecciona al azar un estudiante y es de una estatura mayor de 180 cm, ¿cuál es la probabilidad de que se haya elegido a una mujer?

8. Se dispone de dos métodos para transmitir un mensaje. El método A transmite correctamente el 70% de los mensajes y el método B el 90%. Un día se elige un método al azar y se transmiten ocho mensajes comprobándose posteriormente que los dos primeros se han transmitido de manera incorrecta; ¿cuál es la probabilidad de que se utilizara el método A? y ¿cuál es la probabilidad de que se utilizara el método B?

Resultados en el Anexo.

CAPÍTULO 2

Muestreo y definición de variables aleatorias discretas y continuas


1
10%
12%
20%
40%
1
2
3
4
5
2
3
1
4
6
5
8%
20%
2%
30%
40%
10%
30%
50%
10%

Después de revisar los conceptos fundamentales de estadística y probabilidad, es importante vincular estos temas con el análisis muestral y sus principales técnicas, ya que constituyen herramientas fundamentales para la elaboración de modelos económicos con enfoque estadístico y econométrico.

2.1 MUESTREO y sus técnicas

Si recordamos lo expuesto en el capítulo anterior, sabemos que todos los elementos de estudio se encuentran contenidos dentro del "conjunto universo" (U), mientras que los subconjuntos incluyen solamente algunos de sus elementos. En lo referente a muestreo, una definición clásica y comúnmente aceptada del tema es *la elección de una parte del total de elementos a través de distintas técnicas estadísticas*, por lo que podemos afirmar que un subconjunto del conjunto universo puede ser generado a partir de técnicas de muestreo. En este sentido, puede afirmarse que un subconjunto del conjunto universo se obtiene a partir de procedimientos de muestreo. Si aún tienes dudas, revisa el video de nociones básicas de muestreo (ver video 10).

Una vez definido el término de muestreo, es necesario mencionar que existen dos tipos: el primero es el muestreo probabilístico (aleatorio) y el segundo es el no probabilístico (no aleatorio).

Al realizar un muestreo no probabilístico, la elección de elementos o personas del conjunto universo se lleva a cabo bajo criterios específicos que dejan de lado al azar, por lo que existe la posibilidad de caer en sesgos de selección, ya que de antemano se buscan características especiales en los individuos u objetos de estudio.

El muestreo probabilístico o aleatorio se realiza bajo métodos que consideran el azar, razón por la cual todos los objetos o personas del conjunto universo tienen la misma probabilidad de ser seleccionadas; esta técnica de muestreo es la más recomendada. Para ello, existen distintos métodos de selección, entre los que destacan:

- Aleatorio simple
- Sistemático
- Estratificado
- Por conglomerados

2.2 DIFERENCIAS entre muestreo probabilístico y no probabilístico

En esta sección se busca ejemplificar de forma clara y sencilla cómo iniciar con el tema de muestreo a partir del concepto de probabilístico, por lo que se sugiere pensar en el total de personas electoras de un país (conjunto universo). A partir de ese grupo se elegirá a algunas personas para preguntarles sobre sus preferencias políticas (subconjunto del conjunto universo, es decir una muestra) y así conocer las preferencias electorales en una jornada de votación cualquiera.

Si se utiliza un muestreo aleatorio todas las personas electoras tienen la misma probabilidad de ser elegidas para el estudio, por lo que al encuestarlas encontraremos resultados muy parecidos en caso de haber preguntado a toda la población. En cambio, si se usa un muestreo no aleatorio el encuestador puede elegir a personas que tengan cierta preferencia por un partido político, por lo que el resultado estará sesgado al encontrar preferencias por dicho partido político, lo cual generaría un resultado que no es representativo de la realidad.

En este punto seguramente surge la duda de por qué no preguntar a toda la población y de esa forma encontrar el resultado exacto. La respuesta radica en la facilidad de preguntar a unas cuantas personas en lugar de preguntarle a toda la población, ya que los costos y tiempos de trabajo se reducen cuando se trabaja con una muestra. Tampoco se debe perder de vista que un buen proceso de muestreo garantiza resultados muy cercanos a los obtenidos en caso de preguntar a toda la población.

A manera de resumen, en la siguiente figura se representa las diferencias entre muestreo probabilístico y muestreo no probabilístico.

Figura 2.1
Diferentes tipos de muestreo.
Fuente: elaboración propia.

Una vez explicadas las ventajas de utilizar muestreo probabilístico sobre el no probabilístico, es necesario describir las distintas técnicas del muestreo aleatorio.

2.3 Muestreo ALEATORIO SIMPLE

El muestreo aleatorio simple es la técnica que nos permite conocer cuántos objetos o personas se debe seleccionar del total de la población, permitiendo así pasar del conjunto completo a un subconjunto representativo.

Sobre el total de la población, es necesario decir que esta puede ser finita o infinita; una población finita es aquella en la que conocemos su tamaño (cantidad total), mientras que en la infinita desconocemos esa cifra.

La fórmula para calcular el tamaño óptimo en el muestreo aleatorio simple, en una población finita o infinita así como los valores del nivel de confianza esperado (Z) se muestran en la siguiente tabla.

Parámetro	Población finita	Población infinita
Media	$n = \frac{Z^2 N \sigma^2}{NE^2 + Z^2\sigma^2}$	$n = \frac{Z^2\sigma^2}{E^2}$
Proporción	$n = \frac{Z^2 Npq}{NE^2 + Z^2 pq}$	$n = \frac{Z^2 pq}{E^2}$

Tabla 2.1 Fórmulas para calcular el tamaño de muestra. Fuente: elaboración propia.

La nomenclatura utilizada es:

n=tamaño de la muestra	*N=tamaño total de la población*	*σ^2=varianza*
p=proporción muestral	*q=1-p*	*Z=nivel de confianza a partir de una distribución normal estándar*

En lo referente al nivel de confianza, la siguiente tabla muestra los correspondientes valores para *Z*:

Tabla 2.2. Valores de Z para distintos niveles de confianza. Fuente: elaboración propia.

Nivel de confianza (%)	Z
99.7	3
99	2.58
98	2.33
96	2.05
95	1.96
90	1.645
80	1.28
50	0.674

2.3.1 Forma alternativa para realizar el cálculo del tamaño de la muestra en el muestreo aleatorio simple

Normalmente, el tamaño de la muestra se calcula considerando un subconjunto de características o parámetros poblacionales de interés. Los factores principales en el cálculo de la muestra son:

- Precisión de las estimaciones (intervalos de confianza)
- Costo
- Tiempo
- Diseño muestral

sin dejar de lado el error de muestreo.

a. Controlar el error absoluto d

Consideremos la siguiente fórmula para controlar el error absoluto d.

$$n = \frac{\left(\frac{ts}{d}\right)^2}{1 + \frac{1}{N}\left(\frac{ts}{d}\right)^2}$$

La forma en cómo se calcula S es con la prueba de rango o "regla de dedo" sugerida por Tukey:

$$\sqrt{S^2} = \frac{y_{max} - y_{min}}{4}$$

Ejemplo

1. 1. Suponga que se desea calcular el tamaño de muestra necesario para estimar el ingreso medio de los hogares en la ciudad K.

Los datos son:

N = 17 698

t = 1.96 (con α =0.05)

d = 500

s = 15 751.95

Utilizando la expresión:

$$n = \frac{\left(\frac{ts}{d}\right)^2}{1 + \frac{1}{N}\left(\frac{ts}{d}\right)^2}$$

Sustituyendo los valores de la ciudad K:

$$n = \frac{\left(\frac{(1.96)(15751.95)}{500}\right)^2}{1 + \frac{1}{17698}\left(\frac{(1.96)(15751.95)}{500}\right)^2} = 3137$$

Utilizando:

$$y_{max} = 50,\ 000,\ y_{min} = 1{,}000$$

De donde:

$$\sqrt{s^2} = \frac{50000 - 1000}{4} = 12250$$

Entonces utilizando la fórmula para controlar el error absoluto *d*:

$$n = \frac{\left(\frac{ts}{d}\right)^2}{1 + \frac{1}{N}\left(\frac{ts}{d}\right)^2}$$

Se sustituyen los valores obtenidos.

$$n = \frac{\left(\frac{(1.96)(12250)}{500}\right)^2}{1 + \frac{1}{17698}\left(\frac{(1.96)(12250)}{500}\right)^2} = 2040$$

b. En la proporción

$$n = \frac{\frac{z^2(p)(q)}{e^2}}{1 + \frac{1}{N}\left[\frac{z^2(p)(q)}{e^2} - 1\right]}$$

Ejemplo

N = 18,880

t = 1.96

d = 0.05

p = 0.3

q = 0.7

Si no se puede suponer normalidad, es decir no se conoce el origen de los datos o se trata de una población infinita, entonces se utiliza la siguiente fórmula:

$$n = \frac{1}{\frac{1}{N} + \frac{d^2}{\lambda s^2}} = \frac{\frac{\lambda s^2}{d^2}}{1 + \frac{\lambda s^2}{N d^2}}$$

la cual está basada en la desigualdad de Chevyshev.

Para un nivel de confianza del 95 %, el valor de λ se calcula mediante la siguiente fórmula:

$$1 - \frac{1}{\lambda^2} = 0.95$$

$$\lambda^2 = \frac{1}{0.05} = 20$$

$$\lambda = \sqrt{20} = 4.472$$

Ejemplo

N = 17 698

d = 500

s = 15 751.95

Sustituyendo en la fórmula:

$$n = \frac{1}{\frac{1}{N} + \frac{d^2}{\lambda s^2}}$$

$$n = \frac{1}{\frac{1}{17698} + \frac{(500)^2}{4.4(15751.95)^2}} = 3503$$

Ejemplo

N = 17 698

t = 1.96 (con α=0.05)

d = 500

s = 12, 250

$$n = \frac{1}{\frac{1}{N} + \frac{d^2}{\lambda s^2}}$$

$$n = \frac{1}{\frac{1}{17698} + \frac{(500)^2}{4.4(12250)^2}} = 2298$$

Utilizando Y_{max}=50,000, $Y_{\min}$=1,000

Donde:

$$\sqrt{S^2} = \frac{50{,}000 - 1000}{4} = 12{,}250$$

Por lo tanto:

$$n = \frac{\left[\frac{(1.96)(12{,}250)}{500}\right]^2}{1 + \frac{1}{17698}\left[\frac{(1.96)(12{,}250)}{500}\right]^2} = 2040$$

c. Con la desigualdad de Chevyshev

$$n = \frac{1}{\frac{(18880-1)(0.05)^2}{(18880)(4.4)(0.3)(0.7)} + \frac{1}{18880}} = 363$$

2.3.2 Determinación del tamaño de la muestra considerando los costos fijos y los costos variables

El cálculo del tamaño de una muestra implica una relación entre diversas variables; para este caso, se consideran los costos fijos y los costos variables. Respecto al costo total, la fórmula utilizada para su cálculo, considerando el costo fijo (C_f) y el costo variable (C_v) es:

$$C_t = C_f + C_v \qquad (1)$$

Como el costo variable depende del número de unidades que forman el tamaño de la muestra se tiene la siguiente expresión matemática:

$$C_t = C_f + nC_v \qquad (2)$$

Donde n representa el tamaño de la muestra y al despejarla se tiene lo siguiente:

$$n = \frac{C_t - C_f}{C_v}$$

En el muestreo aleatorio simple el tamaño de la muestra se calcula por medio de:

$$n = \frac{z^2 s^2}{d^2}$$

Donde:

Z es el desvío normal determinado por el nivel de confianza, d es el semi ancho del intervalo de confianza, y S^2 representa la varianza de la población.

Al igualar la ecuación 1 con la ecuación 2, se llega a:

$$\frac{C_t - C_f}{C_v} = \frac{z^2 s^2}{d^2} \qquad (3)$$

La ecuación (3) se despeja para obtener el valor de d:

$$d = zs\sqrt{\frac{C_v}{C_t - C_f}}$$

Ejercicio

Se tiene un presupuesto de $5,000.00 para un estudio de mercado, donde los costos fijos son de $2,000.00 y el costo medio variable es de $15.00. La desviación estándar es de $350.00, un semi ancho de 9 con un nivel de confianza del 90%.

Determinar el tamaño de la muestra y el valor del semi ancho.

Solución:

Datos:

$Ct = \$5000$ $s = \$350$ $Cf = \$2000$ $Cv = \$15$ $d = 9$ $z = 1.64$

El tamaño de la muestra basándose en los costos se determina con la siguiente expresión:

$$n = \frac{C_t - C_f}{C_v}$$

Sustituyendo los valores respectivos:

$$n = \frac{5000 - 2000}{15} = 200$$

Considerando los valores de z, s y de d entonces se utiliza la fórmula:

$$n = \frac{z^2 s^2}{d^2}$$

Sustituyendo:

$$n = \frac{(1.64)^2(350)^2}{(9)^2} = 4067.6049$$

El cálculo del valor de d se realiza por medio de:

$$d = zs\sqrt{\frac{C_v}{C_t - C_f}}$$

Sustituyendo los valores respectivos:

$$d = (1.64)(350)\sqrt{\frac{15}{5000 - 2000}} = 40.58$$

¿Cuál es el presupuesto verdadero para obtener una muestra de 4067.6049? Se utiliza la siguiente expresión:

$$C_t = \frac{C_v z^2 s^2}{d^2} + C_f$$

Sustituyendo:

$$C_t = \frac{(15)(1.64)^2(350)^2}{(9)^2} + 2000$$

$$C_t = 63014.07407$$

La cantidad de 63014.07407 se aplica en la fórmula siguiente:

$$n = \frac{C_t - C_f}{C_v}$$

Para obtener el tamaño de la muestra.

$$n = \frac{63014.07407 - 2000}{15} = 4067.6049$$

El tamaño de la muestra (**4067.6049**) multiplicado por el costo de cada unidad (**15**) da un total de **$61014.07407** y sumándole los costos fijos de **$2000** da la cantidad de **$63014.07407.**

2.4 Muestreo SISTEMÁTICO

En algunos casos, en especial cuando hay grandes poblaciones (imaginemos a la población total de China o India), es tardada la selección de una muestra aleatoria simple cuando se determina primero un número aleatorio y después se cuenta o se busca en la lista de elementos de la población hasta encontrar el elemento correspondiente.

Una alternativa al muestreo aleatorio simple es el muestreo sistemático, que se define como el método para elegir una muestra seleccionando al azar uno de los primeros k elementos y a continuación cada k-ésimo elemento.

Más que un método de muestreo, esta técnica podría ser considerada como un proceso de selección que algunos autores denominan selección a intervalos regulares.

El proceso de selección es el siguiente:

1. Supongamos que una población estudiada está compuesta por 360 elementos, además el tamaño de la muestra es de 30; con esta información se puede determinar el intervalo de selección, simbolizada por I.

$$I = \frac{1}{f} = \frac{1}{\frac{n}{N}} = \frac{N}{n}$$

$$\Rightarrow$$

$$I = \frac{360}{30} = 12$$

2. Hemos determinado el intervalo de selección, ahora se debe obtener un número aleatorio dentro de ese intervalo; supongamos que entre 001 y 012, se obtuvo el número 004, al cual se le denomina punto de arranque.

3. Una vez que se ha establecido el punto de arranque, mediante la selección aleatoria se inicia el proceso de selección sistemática, correspondiendo el número 016 a la segunda unidad seleccionada (004 + 012 = 016); a esta se le suma nuevamente el valor del intervalo, para obtener la tercera unidad y así sucesivamente, dando como resultado los siguiente valores:

004 016 028 040 052 064 076 088 100 112 124 136

148 160 172 184 196 208 220 232 244 256 268 280

292 304 316 328 340 352

4. En el caso de que el valor obtenido para el intervalo de selección no sea un número entero, se procederá de manera diferente. Siendo la población de 355 elementos y el tamaño 30, se tendrá un intervalo igual a:

$$I = \frac{1}{f} = \frac{1}{\frac{n}{N}} = \frac{N}{n}$$

Por lo tanto, el intervalo es:

$$I = \frac{355}{30} = 11.83 \cong 12$$

Considere el siguiente ejemplo: mediante la tabla de números al azar se selecciona un número, entre 001 y 355; supongamos que fue el 208, a partir del cual (punto de arranque) se va acumulando el valor del intervalo, así:

208	220	232	244	256	268	280	292	304	316
328	340	352	009	021	033	045	057	069	081
093	105	117	129	141	153	165	177	189	201

El proceso más utilizado en la práctica es el de calcular el intervalo de selección sin importar si es un número entero o si presenta decimales, en este último caso, se aproxima al número inmediatamente superior, luego se selecciona aleatoriamente (al azar) un número dentro del intervalo, con la finalidad de obtener el punto de arranque, al cual se le va acumulando el valor del intervalo.

2.5 Muestreo ESTRATIFICADO

Este tipo de muestreo sirve para conocer un tamaño de muestra y después colocar una parte de esa muestra en distintos estratos, los cuales deben de ser internamente homogéneos, de tal forma que justifiquen y optimicen el costo de la estratificación. Por ejemplo, para un estudio de preferencias electorales puede resultar interesante conocer la opinión de los habitantes de la zona sur y norte del país.

Si la población de la zona norte es el 40% respecto al total del país y la del sur representa el 60%, se tomará una muestra que contenga también esa misma proporción.

Dentro del muestreo estratificado, existen distintas técnicas, las cuales son:

- Por asignación proporcional.
- Por asignación óptima.
- Por asignación óptima económica.

A continuación, se desarrollarán cada una de las técnicas del muestreo estratificado.

2.5.1 Por asignación proporcional

Este muestreo estratificado reparte el tamaño de la muestra en forma proporcional al tamaño del estrato.

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

El tamaño de la muestra en cada estrato es:

$$n_i = n(wp_i)$$

Se calculan tantas W_{pi} como estratos se hayan definido con las restricciones de que:

$$\sum_{i=1}^{k} wp_i = 1$$

$$\sum_{i=1}^{k} n_i = n$$

Ejemplo

En unos sectores empresariales se encontró que de los $7{,}500$ clientes inscritos en programas de capacitación, su distribución se presentó de la siguiente manera:

Tabla 2.3
Distribución de sectores empresariales.
Fuente: elaboración propia.

Sector	Población
Uno	1,700
Dos	2,500
Tres	2,000
Cuatro	1,300
Total	**7,500**

Se desea asignar de manera proporcional una muestra de 50 unidades entre los 4 estratos.

Aplicando la expresión:

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

Entonces:

$$wp_1 = \frac{1700}{7500} = 0.226662$$

$$wp_2 = \frac{2500}{7500} = 0.333333$$

$$wp_3 = \frac{2000}{7500} = 0.266666$$

$$wp_4 = \frac{1300}{7500} = 0.173333$$

Para verificar que se está considerando todos los estratos, comprobamos que la sumatoria resulte en la unidad.

$$\sum_{i=1}^{4} wp_i = 0.226662 + 0.333333 + 0.266666 + 0.173333 = 1$$

Finalmente, se calcula el tamaño de la muestra para cada estrato.

$$n_1 = (0.226662)50 = 11.333333 \approx 11$$

$$n_2 = (0.019905727)50 = 0.99528635 \approx 1$$

$$n_3 = (0.266666)50 = 13.333333 \approx 13$$

$$n_4 = (0.173333)50 = 8.666666 \approx 9$$

Y se comprueba que la sumatoria sea igual al tamaño de la muestra establecida, en este caso 50.

$$\sum_{i=1}^{k} n_i = 11 + 17 + 13 + 9 = 50$$

2.5.2 Por asignación óptima

La asignación óptima considera la proporción de la población y la variabilidad o dispersión de los datos. Es útil cuando los costos de muestreo por unidad para cada estrato son similares o no importan y cuando las varianzas son diferentes en cada estrato. Cuando más grande sea una varianza, mayor tamaño de muestra requerirá ese estrato.

La expresión matemática es:

$$wo_i = \frac{N_i s_i}{\sum_{i=1}^{k} N_i s_i}$$

Para el caso de proporciones:

$$s = \sqrt{\bar{P}(1-\bar{P})}$$

El tamaño de la muestra en cada estrato es:

$$n_i = n(wo_i)$$

Con las restricciones de:

$$\sum_{i=1}^{k} (wo_i) = 1$$

Además:

$$\sum_{i=1}^{k} n_i = n$$

Ejemplo

Ahora se desea asignar de manera óptima una muestra de 50 unidades entre los 4 estratos, conociendo que las desviaciones estándar estimadas del monto de las ventas por estrato son:

$S_1 = 273\ 500$

$S_2 = 5\ 870$

$S_3 = 28\ 700$

$S_4 = 154\ 000$

La justificación para usar la asignación óptima consiste en la diferencia entre las desviaciones estándar estimadas para cada estrato.

Para el cálculo de la sumatoria de Ni Si, tenemos:

Tabla 2.4
Cálculo de la sumatoria NiSi. Fuente: elaboración propia.

Ni	Si	NiSi
1 700	273 500	464 950 000
2 500	5 870	14 675 000
2 000	28 700	57 400 000
1 300	154 000	200 200 000
		Σ = 737 225 000

Aplicamos la expresión matemática:

$$wp_1 = \frac{464950000}{737225000} = 0.630675845$$

$$wp_2 = \frac{14675000}{737225000} = 0.019905727$$

$$wp_3 = \frac{57400000}{737225000} = 0.07785954$$

$$wp_4 = \frac{200200000}{737225000} = 0.271558886$$

Ahora calculamos el tamaño de la muestra para cada estrato:

$$n_1 = (0.630675843)50 = 31.53379215 \approx 31$$
$$n_2 = (0.019905727)50 = 0.99528635 \approx 1$$
$$n_3 = (0.07785954)50 = 3.892977 \approx 4$$
$$n_4 = (0.271558886)50 = 13.5779443 \approx 14$$

Para finalmente comprobar que nuestro tamaño de muestra es igual al establecido en el planteamiento del problema.

$$\sum_{i=1}^{k} n_i = 31 + 1 + 4 + 14 = 50$$

2.5.3 Por asignación óptima económica

Para este tipo de muestreo, considera la proporción de la población, la variabilidad o dispersión de los datos y el costo de muestreo por unidad en cada estrato.

Para el caso de proporciones, la fórmula es:

$$s = \sqrt{\bar{P}(1-\bar{P})}$$

El tamaño de la muestra en cada estrato se calcula mediante la expresión:

$$n_i = n(wct_i)$$

Se calculan tantas wct_i como estratos sean con las siguientes restricciones:

$$\sum_{i=1}^{k}(wct_i) = 1$$

y

$$\sum_{i=1}^{k} n_i = n$$

Ejemplo

Utilizaremos nuevamente el enfoque planteado en la asignación proporcional para llevar a cabo el desarrollo correspondiente y puedan contrastarse los resultados.

De un total de $7,500$ clientes de diversos sectores empresariales que se inscribieron para capacitación, sabemos que $1{,}700$ pertenecen al sector 1; $2{,}500$ al sector 2; $2{,}000$ al sector 3 y $1{,}300$ al sector 4.

Se desea asignar de manera *óptima económica* una muestra de 50 unidades. Conociendo que las desviaciones estándar estimadas por estrato son de:

$s_1 = 273\ 500$

$s_2 = 5\ 870$

$s_3 = 28\ 700$

$s_4 = 154\ 000$

y los costos de muestreo por unidad son de:

$C_1 = \$2\ 500$

$C_2 = \$\ 900$

$C_3 = \$1\ 100$

$C_4 = \$1\ 200$

Como se puede observar, se hace evidente una amplia diferencia de costos de muestreo por estrato lo cual justifica el uso de esta técnica para obtener resultados más precisos; esto nos lleva a la siguiente tabla:

Tabla 2.5
Ejemplo por asignación óptima económica.
Fuente: elaboración propia.

Ni	si	Nisi	C_i	$\sqrt{C_i}$	$\frac{N_i s_i}{\sqrt{C_i}}$
1 700	273 500	464 950 000	2 500	50	9 299 000
2 500	5 870	14 675 000	900	30	489 166.667
2 000	28 700	57 400 000	1 100	33.17	1 730 479.349
1 300	154 000	200 200 000	1 200	34.64	5 779 445.727
					Σ=17 298 091.743

Para el cálculo de la sumatoria de $\frac{N_i s_i}{\sqrt{C_i}}$ se realizan las siguientes operaciones:

$$wct_1 = \frac{9299000}{17298091.743} = 0.53757 \approx 0.54$$

$$wct_2 = \frac{489166.6667}{17298091.743} = 0.02827 \approx 0.03$$

$$wct_3 = \frac{1730479.349}{17298091.743} = 0.1000 = 0.10$$

$$wct_4 = \frac{5779445.727}{17298091.743} = 0.3341 \approx 0.33$$

Lo que nos da como resultado final:

$$\sum_{i=1}^{4} wct_i = 0.54 + 0.03 + 0.10 + 0.33 = 1$$

Por lo tanto, el tamaño de la muestra para cada estrato es:

$$n_1 = 50(0.54) = 27$$

$$n_2 = 50(0.03) = 1.5 = 2$$

$$n_3 = 50(0.10) = 5$$

$$n_4 = 50(0.33) = 16.5 = 16$$

Lo que nos da como resultado final

$$\sum_{i=1}^{4} n_i = 27 + 2 + 5 + 16 = 50$$

2.6 Muestreo por CONGLOMERADOS

Los métodos presentados hasta ahora están pensados para seleccionar directamente los elementos de la población, es decir, que las unidades muestrales forman parte de los elementos de la población.

En el muestreo por conglomerados la unidad muestral es un grupo de elementos de la población que forman una unidad, a la que llamamos conglomerado.

Las unidades hospitalarias, los departamentos universitarios, una selección de determinados productos o de urnas electorales, por citar algunos casos, son conglomerados naturales.

Cuando los conglomerados son áreas geográficas, suele hablarse de “muestreo por áreas”.

El muestreo por conglomerados consiste en seleccionar aleatoriamente un cierto número de conglomerados (el necesario para alcanzar el tamaño muestral establecido) y posteriormente analizar todos los elementos pertenecientes a los conglomerados elegidos.

La metodología consiste en dividir a la población en conjuntos separados de elementos, llamados conglomerados; cada elemento de la población pertenece a un solo grupo.

Después se toma una muestra aleatoria simple de los conglomerados. Todos los elementos dentro de cada conglomerado muestreado forman parte de la muestra.

El muestreo por conglomerados tiende a proporcionar los mejores resultados cuando los elementos de los conglomerados son heterogéneos (desiguales).

En el caso ideal, cada conglomerado es una versión representativa, en pequeña escala, de toda la población. La eficacia del muestreo por conglomerados depende del grado de representatividad que cada conglomerado tiene con respecto al total de la población.

Esta técnica tiene utilidad cuando el universo que se requiere estudiar admite ser subdividido en universos menores de características similares a las del universo total y se procede a subdividir el universo en un número finito de conglomerados y, entre ellos, se pasa a elegir algunos que serán los únicos que se investigarán; esta elección puede realizarse por el método del muestreo simple o por el del muestreo sistemático.

Una vez cumplida esta etapa, puede efectuarse una segunda selección, dentro de cada uno de los conglomerados elegidos, para llegar a un número aún más reducido de unidades muestrales.

La ventaja de esta técnica es que obvia la tarea de confeccionar el listado de todas las unidades del universo. Su desventaja mayor radica en que, al efectuarse el muestreo en dos etapas los errores muestrales de cada una se van acumulando, lo que da un error mayor que en cualquiera de los métodos anteriores.

La técnica de conglomerados suele utilizarse cuando queremos extraer muestras de los habitantes de un conjunto geográfico amplio, por ejemplo una gran ciudad o un conjunto de pueblos, por lo que se procede a tomar cada pueblo o grupo de manzanas como un conglomerado independiente; otros ejemplo de uso son para conocer las reservas forestales y marinas, para estudiar las estrellas u otros casos semejantes.

Ejemplo

En México existen 32 entidades federativas, las cuales tienen distintas actividades productivas. Si se desea conocer la productividad y salarios del país, una opción es hacer el cálculo a partir de la totalidad de entidades, la otra, seleccionar solo algunas entidades, las cuales pueden ser las más activas (Nuevo León, Jalisco, Querétaro, Guanajuato y Aguascalientes).

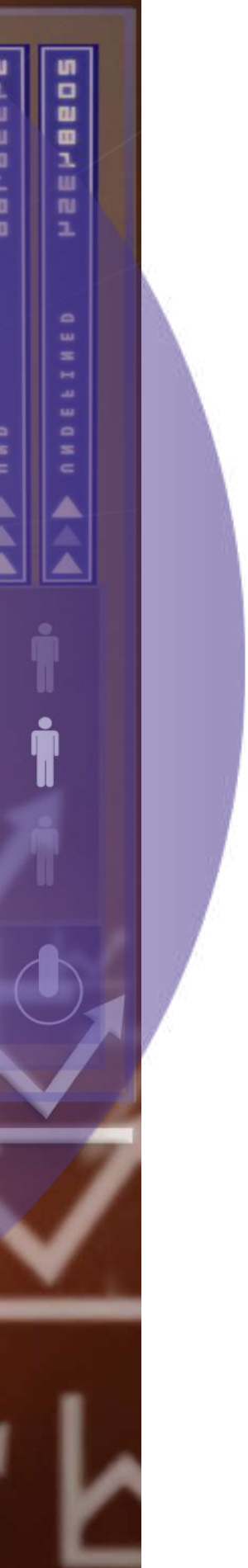
Una vez hecho los conglomerados (entidades seleccionadas), se realiza el análisis en cada una de ellas. Para cerrar esta sección, el siguiente video muestra la forma de cómo pasar del total de la población a un tamaño de muestra usando como ejemplo a la Población Económicamente Activa de México (ver video 11).

https://youtu.be/vj_Tcu8fWNI

CAPÍTULO 3

Definición de variable aleatoria discreta, aleatoria continua y sus principales funciones de distribución


95%
80%
75%
46%
21%
35%



Para poder establecer de forma clara cuáles son las diferencias entre una variable aleatoria discreta y una variable aleatoria continua es importante comenzar definiendo, en lo general, lo que es una variable aleatoria.

Una variable aleatoria se puede definir como una función que solo puede tomar números reales asociados a los resultados posibles del espacio muestral del experimento, es decir:

$$X : \Omega \rightarrow \mathbb{R}$$

Por lo tanto, los posibles resultados de dicho experimento aleatorio son los n números reales de x que la función X puede tomar.

En la figura 1 vemos una representación gráfica de una variable aleatoria como función de Ω en los $\mathbb{R}$.

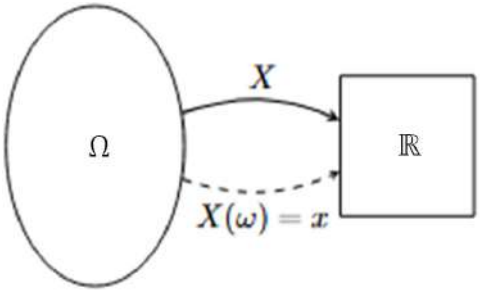


Figura 3.1 Representación de una variable aleatoria como función desde Ω hacia $\mathbb{R}$. Fuente: elaboración propia.

Para ejemplificar lo anterior, pensamos en lanzar un dado y buscar la probabilidad de obtener un número par:

$$v = \{ 1, 2, 3, 4, 5, 6 \} \qquad (1)$$

$$X = \{ 2, 4, 6 \} \qquad (2)$$

En el espacio muestral (1) se encuentra el conjunto universo, es decir, todas las caras del dado. En el subconjunto (2) se indica el espacio muestral que cumple con la condición de obtener un número par al realizar el lanzamiento del dado.

Por tanto, el subconjunto X es la variable aleatoria, ya que es ahí donde se encuentran los valores que hacen referencia a los números pares después de lanzar un dado.

Es importante mencionar en este punto que cuando se coloca la X en mayúscula, se hace referencia a la función de la variable aleatoria, mientras que cuando se hace en minúsculas, nos referimos únicamente a los valores que puede tomar la función.

Una vez definido el concepto de variable aleatoria es necesario hacer la diferenciación entre variable aleatoria discreta y variable aleatoria continua, para así dar paso a las funciones de densidad y de distribución (ver video 12).

https://youtu.be/1SVv_yhtnLE

3.1 Variables ALEATORIAS DISCRETAS

Una variable aleatoria se considera discreta cuando toma valores dentro de un conjunto numerable y su comportamiento se describe mediante una función de probabilidad.

Por lo tanto, decimos que:

$$P(X = X_k) = f(x_k) \qquad \forall k \in \{1,2,3 \dots n)$$

Dicho de otro modo, **una variable aleatoria discreta sólo puede tomar valores enteros**, lo cual limita su uso a ciertos casos. Por ejemplo, pensemos en la correcta elaboración de un bien manufacturado, el lanzamiento de una moneda o de un dado.

En todos los casos mencionados estamos hablando de un espacio muestral que hace referencia a números enteros; esto se debe a que un bien solo existe cuando ha sido manufacturado completamente, o bien, al lanzar una moneda o un dado, únicamente es posible obtener un resultado definido. Por lo tanto, estos escenarios son analizados desde la óptica de variables aleatorias discretas.

Para ejemplificar lo anterior, en la siguiente figura observamos gráficamente el comportamiento de una variable aleatoria discreta, para un experimento de lanzamiento de moneda.

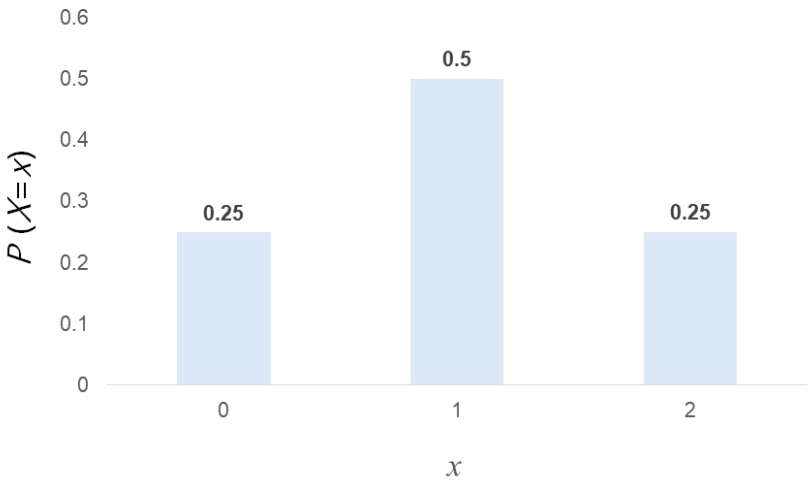


Figura 3.2 Distribución discreta: número de caras en 2 lanzamientos de moneda.
Fuente: elaboración propia a partir del procesamiento de datos.

Como habíamos mencionado anteriormente, el usar la letra x en minúscula hace referencia a los valores que puede tomar la función. Por tanto, en este caso, la línea horizontal hace referencia a los valores que puede tomar la probabilidad de obtener una cara al lanzar dos monedas (0, 1 y 2), es decir:

- Probabilidad de encontrar cero caras al lanzar dos monedas: 25%
- Probabilidad de encontrar una cara al lanzar dos monedas: 50%
- Probabilidad de encontrar dos caras al lanzar dos monedas: 25%

Como es una variable discreta, tiene valores específicos y se representa mediante una gráfica de barras, nunca una curva continua.

3.2 Variables ALEATORIAS CONTINUAS

Una variable aleatoria es continua si toma un número infinito de valores dentro de un intervalo; los números enteros y decimales están considerados dentro del espacio muestral.

En la sección anterior, mencionamos que la elaboración de artículos en la manufactura puede ser abordado desde la óptica de una variable aleatoria discreta, ya que se trata de números enteros.

Si en lugar de bienes manufacturados pensamos en describir los ingresos de la población ocupada, estamos ante la posibilidad de utilizar variables aleatorias continuas, ya que los salarios son pagados en pesos y centavos, lo cual implica usar números decimales.

El video 13 muestra cómo es que las variables aleatorias continuas pueden ser incorporadas a la economía a partir de distintas funciones y ecuaciones como la de Kuznets, Phillips y Kuznets ambiental.

3.3 Funciones de distribución de probabilidad de las variables ALEATORIAS DISCRETAS

Antes de iniciar con el estudio de las funciones de distribución de probabilidad de variables aleatorias discretas, es necesario conocer las condiciones generales que se deben cumplir.

Consideremos a X una variable aleatoria discreta con valores x_0, x_1, ... y su función de probabilidad denotada como $R \rightarrow R$, la cual se define como:

$$f(x) = \begin{cases} \mathbb{P}(X = x) & si\ x = x_1, x_2, \ldots\ldots \\ 0 & otro\ caso. \end{cases}$$

Por lo tanto, es una función $f(x) = P(X = x)$ tal que, cumple:

$$0 < f(x) \leq 1$$

$$\sum_x f(x)n = 1$$

Una vez definidas las condiciones, en la siguiente tabla se muestran las principales funciones discretas de probabilidad para las variables aleatorias discretas.

Distribución	Función de probabilidad	Parámetros	Esperanza	Varianza
Uniforme discreta	$f(x) = \frac{1}{n}$	$x_1, x_2, \ldots n \in \mathbb{R}$	$\mathbb{E}(X) = \frac{1}{n}\sum_{i=1}^{n} x_i = \mu$	$\mathbb{V}ar(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$
Bernoulli	$\mathbb{P}(X = x) = f(x) = p^x(1-p)^{1-x}$	$0 < p < 1$	$\mathbb{E}(X) = p$	$\mathbb{V}ar(X) = p(1-p)$
Binomial	$f(x) = \mathbb{P}(X = x) = \binom{n}{x} p^x(1-p)^{n-x}$	$n = 1,2, \ldots$ $0 < p < 1$	$\mathbb{E}(X) = np$	$\mathbb{V}ar(X) = np(1-p)$
Geométrica	$f(x) = \mathbb{P}(X = x) = (1-p)^x P$	$0 < p < 1$	$\mathbb{E}(X) = \frac{1-p}{p}$	$\mathbb{V}ar(X) = \frac{1-p}{p^2}$
Binomial Negativa	$f(x) = \mathbb{P}(X = x) = \binom{r+x-1}{x} p^r(1-p)^x$	$r = 1,2, \ldots$ $0 < p < 1$	$\mathbb{E}(X) = \frac{r(1-p)}{p}$	$\mathbb{V}ar(X) = \frac{r(1-p)}{p^2}$
Hipergeométrica	$f(x) = \mathbb{P}(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$	$k - n = 1.2. \ldots$	$\mathbb{E}(X) = \frac{rn}{N}$	$\mathbb{V}ar(X) = \frac{rn}{N}\left[\frac{(n-1)(r-1)}{N-1} + 1 - \frac{rn}{N}\right]$
Poisson	$f(x) = \mathbb{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$	$\lambda > 0$	$\mathbb{E}(X) = \lambda$	$\mathbb{V}ar(X) = \lambda$
Logaritmica	$f(x) = \mathbb{P}(X = x) = -\frac{1}{log(1-p)}\frac{p^x}{x}$	$k = 1,2, \ldots$	$\mathbb{E}(X) = \frac{ap}{log(1-p)}$	$\mathbb{V}ar(X) = \mathbb{E}(X)\left(\frac{1}{1-p} - \mathbb{E}(X)\right)$

Tabla 3.1 Funciones de distribución discretas. Fuente: elaboración propia.

3.3.1 Función de distribución de probabilidad uniforme discreta

La distribución uniforme discreta es una distribución de probabilidad simétrica que surge en espacios de probabilidad equiprobables; es decir, en situaciones donde, de n rresultados posibles, todos tienen la misma probabilidad de ocurrir.

Su función de probabilidad está dada por:

$$f(x) = \mathbb{P}(X = x) = \frac{1}{n}$$

Para calcular la esperanza y la varianza de esta distribución se utilizan las siguientes fórmulas, respectivamente:

$$\mathbb{E}(X) = \frac{1}{n}\sum_{i=1}^{n} x_i = \mu$$

$$\mathbb{V}ar(X) = \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^2$$

Existe un par de ecuaciones en particular para calcular dichos valores y son las siguientes:

$$\mathbb{E} = \frac{x_f + x_i}{2} \qquad \mathbb{V} = \frac{(x_f - x_i + 1)^2 - 1}{12}$$

Tenemos que mencionar que estas fórmulas sólo pueden ser ocupadas para el caso en que los datos sean enteros y consecutivos.

Ejemplo

Suponga que se lanza un dado (6 caras) donde X es la variable que representa el número que saldrá de dicho experimento

¿Cuál es la probabilidad de que salga algún número?

Decimos que los números son los siguientes $\{1, 2, 3, 4, 5, 6\}$; por lo tanto decimos que la probabilidad determinada es de $\frac{1}{6}$.

En la siguiente tabla se muestra cómo se distribuye la probabilidad para cada número.

Tabla 3.2 Probabilidades para cada número. Fuente: elaboración propia.

Valor	Probabilidad
$X = 1$	1/6
$X = 2$	1/6
$X = 3$	1/6
$X = 4$	1/6
$X = 5$	1/6
$X = 6$	1/6

Por lo tanto, decimos que:

$$f(x) = \begin{cases} \frac{1}{6} & x = x_1, x_2, \dots x_n \\ 0 & en\ cualquer\ otro\ caso \end{cases}$$

Ahora, para calcular la esperanza matemática y la varianza usaremos las ecuaciones anteriormente denotadas:

$$\mathbb{E} = \frac{6+1}{2} \approx 3.5 \qquad \mathbb{V} = \frac{(6-1+1)^2 - 1}{12} \approx 3.083$$

3.3.2 Distribución Bernoulli

La distribución de Bernoulli representa la forma más sencilla de una distribución de probabilidad discreta, y capta los resultados de experimentos que solo pueden dar lugar a dos resultados posibles típicamente etiquetados como éxito o fracaso.

Un ensayo Bernoulli se define como aquel experimento aleatorio con sólo dos resultados posibles:

F = Fracaso (también se denota con la letra “q”).

E = Éxito (también se denota con la letra “p”).

Por tanto, se define la variable aleatoria X como aquella que lleva el resultado de éxito y se asocia al número 1 y el resultado de fracaso al número 0, por lo tanto, decimos que:

$$P \in (0,1)$$

Y su función de densidad de probabilidad está denotada como:

$$\mathbb{P}(X = x) = f(x) = p^x(1-p)^{1-x}$$

Para calcular la esperanza matemática y la varianza usamos las siguientes ecuaciones:

$$\mathbb{E}(X) = p$$

$$\mathbb{V}ar(X) = p(1-p)$$

Ejemplo

En la fabricación de neumáticos se seleccionan de manera aleatoria, tres de ellos; se hace una inspección de los neumáticos y se clasifican en defectuosos y no defectuosos. El proceso de fabricación produce en total el 20% de neumáticos defectuosos. Se considera un éxito la obtención de un artículo defectuoso.

A continuación, se presenta el espacio muestral después de seleccionar tres neumáticos al azar, en el cual D es un neumático defectuoso y ND es un neumático no defectuoso.

Tabla 3.3
Ejemplo de distribución de Bernoulli.
Fuente: elaboración propia.

Resultado	x
(ND)(ND)(ND)	0
(**D**) (ND)(ND)	1
(ND) (**D**) (ND)	1
(ND)(ND) (**D**)	1
(ND) (**D**) (**D**)	2
(**D**) (ND) (**D**)	2
(**D**) (**D**) (ND)	2
(**D**) (**D**) (**D**)	3

Por tanto:

p = 0.20 q=0.80

Se calculan las probabilidades respectivas:

P[(ND)(ND)(ND)] = P(ND)P(ND)P(ND) = (0.80) (0.80) (0.80) = 0.512

P[(D)(ND)(ND)] = P(D)P(ND)P(ND) = (0.20) (0.80) (0.80) = 0.128

P[(ND)(D)(ND)] = P(ND)P(D)P(ND) = (0.80) (0.20) (0.80) = 0.128

P[(ND)(ND)(D)] = P(ND)P(ND)P(D) = (0.80)(0.80)(0.20) = 0.128

$\sum = 0.384$

$$P[(ND)(D)(D)] = P(ND)P(D)P(D) = (0.80)(0.20)(0.20) = 0.032$$

$$P[(D)(ND)(D)] = P(D)P(ND)P(D) = (0.20)(0.80)(0.20) = 0.032$$

$$P[(D)(D)(ND)] = P(D)P(D)P(ND) = (0.20)(0.20)(0.80) = 0.032$$

$$\sum = 0.096$$

$$P[(DDD)] = P(D)P(D)P(D) = (0.20)(0.20)(0.20) = 0.008$$

Con los cálculos anteriores se obtiene la distribución de probabilidad de x:

x	0	1	2	3
$f(x)$	0.512	0.384	0.096	0.008

Por lo tanto, se obtienen los siguientes resultados:

- Cuando no se tienen neumáticos defectuosos la probabilidad es de: 0.512.
- Cuando se tiene un neumático defectuoso la probabilidad es de: 0.384.
- Cuando se tienen dos neumáticos defectuosos la probabilidad es de: 0.096.
- Cuando se tienen tres neumáticos defectuosos la probabilidad es de: 0.008.

El número de éxitos en n experimentos de Bernoulli recibe el nombre de variable aleatoria binomial, la cual será la siguiente distribución por analizar.

3.3.3 Distribución binomial

La distribución binomial es una forma de calcular cómo se distribuyen las probabilidades en situaciones donde solamente hay dos posibles resultados: un éxito o un fracaso.

Se aplica cuando realizamos una serie de pruebas, todas independientes entre sí y queremos saber cuántas veces ocurrirá un evento específico por ejemplo, obtener cara al lanzar una moneda varias veces.

Imaginemos que tenemos n ensayos Bernoulli independientes cada uno del otro, siempre con la misma probabilidad de éxito o fracaso:

$$p \in (0,1)$$

Se dice que una variable aleatoria discreta X tiene distribución binomial con parámetros $n \in \mathbb{N}^+$, por lo que su función de densidad de probabilidad está denotada como:

$$f(x) = \mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Donde $\binom{n}{x}$ es la siguiente expresión:

$$\frac{n!}{x!\,(n-x)!}$$

Para calcular la esperanza matemática y la varianza usamos las siguientes ecuaciones

$$\mathbb{E}(X) = np$$

$$\mathbb{V}ar(X) = np(1-p)$$

Ejemplo

Imaginemos que un 80% de personas en el mundo ha visto el partido de la final del último mundial de fútbol. Tras el evento, 4 amigos se reúnen a conversar, ¿cuál es la probabilidad de que 3 de ellos hayan visto el partido?

$$\mathbb{P}_{(3)} = \frac{4!}{3!\,(4-3)!} 0.8^3 0.2^{(4-3)} = 0.4096 \approx 40.96\%$$

3.3.4 Distribución Poisson

La distribución Poisson es una distribución de probabilidad discreta que expresa, a partir de una frecuencia de ocurrencia media, la probabilidad de que ocurra un determinado número de eventos durante cierto periodo de tiempo. Concretamente, se especializa en la probabilidad de ocurrencia de sucesos con probabilidades muy pequeñas, o sucesos «raros».

Una variable aleatoria discreta tiene distribución Poisson con un parámetro λ >0, y su función de densidad de probabilidad está denotada por:

$$f(x) = \mathbb{P}(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad \forall\, x = \{0,1,2,\dots,n\}$$

Para calcular la esperanza matemática y la varianza usamos las siguientes ecuaciones.

$$\mathbb{E}(X) = \lambda$$

$$\mathbb{V}ar(X) = \lambda$$

Ejemplo

Suponemos que estamos en temporada de invierno y queremos ir a esquiar antes de diciembre. La probabilidad que abran las estaciones de esquí antes de diciembre es del 5%. De las 100 estaciones de esquí que existen queremos saber la probabilidad de que la más cercana abra antes de diciembre. La valoración de dicha estación de esquí es de 6 puntos.

$$\mathbb{P}(X = 6) = \frac{e^{(-5)}5^6}{6!} = 0.1462$$

Por lo tanto, la probabilidad de que la estación de esquí más cercana abra antes de diciembre es de 14.62%.

3.3.5 Distribución geométrica

La distribución geométrica es la distribución de probabilidad del número de fracasos obtenidos hasta conseguir el primer éxito en un proceso de Bernoulli. Una variable aleatoria discreta tiene distribución geométrica con probabilidad $p \in (0,1)$; y su función de densidad de probabilidad está denotada como:

$$f(x) = \mathbb{P}(X = x) = (1-p)^x P$$

Para calcular la esperanza matemática y la varianza usamos las siguientes ecuaciones:

$$\mathbb{E}(X) = \frac{1-p}{p}$$

$$\mathbb{V}ar(X) = \frac{1-p}{p^2}$$

Ejemplo

La probabilidad de que un foco sea producido de manera defectuosa es $p = 0.2$. ¿Cuál es la probabilidad de que el primer foco defectuoso se encuentre en el tercer intento?

$$\mathbb{P}(X = 3) = (1-0.2)^3(0.2)$$

$$= (0.8)^3(0.2) = 0.512(0.2)$$

$$= 0.1024$$

3.3.6 Distribución binomial negativa

La distribución binomial negativa modela el número de ensayos necesarios hasta obtener r éxitos. Esta distribución es una generalización de la distribución geométrica (que es el caso cuando $r = 1$).

Una variable aleatoria discreta tiene distribución binomial negativa con parámetros $r \in \mathbb{R}$ y , $p \in (0,1)$ si su función de densidad de probabilidad está denotada como:

$$f(x) = \mathbb{P}(X = x) = \binom{r + x - 1}{x} p^r (1-p)^x$$

Para calcular la esperanza matemática y la varianza usamos las siguientes ecuaciones:

$$\mathbb{E}(X) = \frac{r(1-p)}{p}$$

$$\mathbb{V}ar(X) = \frac{r(1-p)}{p^2}$$

Ejemplo

Considera que lanzas un dado y la probabilidad de sacar un 6 es $p = \frac{1}{6}$ ¿Cuál es la probabilidad de que necesites 5 lanzamientos para obtener 2 seises?

$$\mathbb{P}(X = 5) = \binom{4}{1}\binom{5}{6}^{(3)}\binom{1}{6}^{(2)} = 0.064$$

3.3.7 Distribución hipergeométrica

La distribución hipergeométrica modela el número de éxitos al sacar una muestra **sin reemplazo** de una población finita que contiene una cantidad conocida de éxitos y fracasos. Se dice que una variable aleatoria discreta tiene distribución hipergeométrica con parámetros $p \in (0,1)$, y con una función de densidad de probabilidad denotada como:

$$f(x) = \mathbb{P}(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

Para calcular la esperanza matemática y la varianza usamos las siguientes ecuaciones:

$$\mathbb{E}(X) = \frac{rn}{N}$$

$$\mathbb{V}ar(X) = \frac{rn}{N}\left[\frac{(n-1)(r-1)}{N-1} + 1 - \frac{rn}{N}\right]$$

Ejemplo

Una caja tiene 10 bolas de colores, de las cuales 4 son rojas y 6 son azules. Se sacan 3 sin reemplazo. ¿Cuál es la probabilidad de sacar exactamente 2 rojas?

Lo primero que se hace es identificar los valores para cada variable:

$$r = 4, \quad N = 10, \quad n = 3, \quad x = 2$$

Para después sustituir en la fómula:

$$\mathbb{P}(X = 2) = \frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}} \approx 0.3$$

3.3.8 Distribución logarítmica

La distribución logarítmica se usa para modelar datos discretos con muchas observaciones pequeñas y pocas observaciones grandes. Se dice que una variable aleatoria discreta tiene distribución logarítmica con parámetros $p \in (0,1)$ y con una función de densidad de probabilidad denotada como:

$$f(x) = \mathbb{P}(X = x) = -\frac{1}{log(1-p)}\frac{p^x}{x}$$

Para calcular la esperanza matemática y la varianza usamos las siguientes ecuaciones:

$$\mathbb{E}(X) = \frac{ap}{log(1-p)}$$

$$\mathbb{V}ar(X) = \frac{ap(1-ap)}{(1-p)^2} = \mathbb{E}(X)\left(\frac{1}{1-p} - \mathbb{E}(X)\right)$$

Donde $a = -\frac{1}{log(1-p)}$

Ejemplo

Si se tiene un valor de $p = 0.5$ ¿Cuál es la probabilidad de que una observación tome el valor $k = 2$?

$$\mathbb{P}(X = 2) = \frac{1}{log(1-0.5)}\frac{0.5^2}{2} = 0.18$$

3.3 Funciones de distribución de probabilidad de las variables ALEATORIAS CONTINUAS

Una variable aleatoria continua puede tomar un número infinito de valores dentro de un intervalo, tal como se ha mencionado, mientras que su función de densidad se define como X siendo una variable aleatoria continua, lo que nos lleva a decir que dicha función integrable $f(x)\colon \mathbb{R} \to [0,\infty)$ es una función de densidad de X sí y sólo sí para cualquier intervalo ($a,\ b$) de $\mathbb{R}$ se cumple:

$$f(x) = \mathbb{P}\big(X \in (a,b)\big) = \int_a^b f(x)dx$$

El valor que tome la variable debe caer en el intervalo (a, b) y también debe cumplir las siguientes propiedades:

$$f(x) \geq 0 \qquad\qquad \forall\, x \in \mathbb{R}$$

$$\int_{-\infty}^{\infty} f(x)\ dx = 1.$$

Toda función $f(x)\colon \mathbb{R} \to [0,\infty)$ que satisfaga estas dos propiedades, sin necesidad de tener una variable aleatoria de por medio se llama **función de densidad** (Rincón, 2017).

En la siguiente figura vemos una representación gráfica de dicha función.

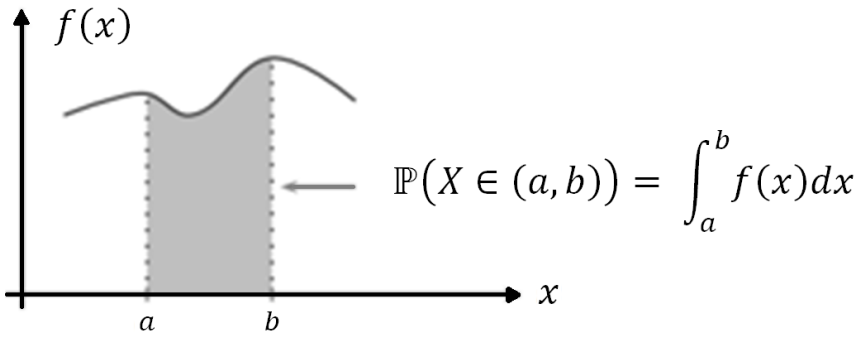


Figura 3.3 Representación gráfica de la función de densidad.
Fuente: elaboración propia a partir del procesamiento de datos.

En lo que se refiere a las funciones de distribución de las variables aleatorias continuas, comenzamos con la de distribución acumulativa.

3.4.1 Función de distribución acumulativa

La distribución acumulada o acumulativa $F(x)$ de una variable aleatoria continua X, con una función de densidad $f(x)$ es:

$$F(x) = \mathbb{P}(X = x) = \int_{-\infty}^{x} f(t)dt \quad \forall \ -\infty \leq x \leq \infty$$

A partir de la definición de función de distribución acumulada de una variable aleatoria continua se deducen las propiedades siguientes:

1. $F(-\infty) = 0$
2. $F(\infty) = 1$
3. $P(x1 \leq X \leq x2) = F(x2) - F(x1)$
4. $\frac{dF(x)}{dx} = f(x)$

3.4.1.1 Esperanza matemática

Sea X una variable aleatoria con función de densidad $f(x)$; se llama esperanza matemática o valor esperado, valor medio o media de X al número real resultante de la siguiente expresión:

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

Significado de la esperanza

Representa una medida de centralización como valor medio teórico de todos los valores que puede tomar la variable ponderados por su probabilidad de ocurrencia.

3.4.1.2 Varianza

Es la medida del cuadrado de la distancia promedio entre la media y cada elemento de la población.

Sea X una variable aleatoria con distribución de probabilidad f(x) y media μ. La varianza de X es calculada por medio de:

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$$

3.4.1.3 Desviación estándar

Es una medida de dispersión de la misma dimensión física de la variable y está indicada por la letra σ.

3.4.2 Distribución uniforme

Para una variable aleatoria que toma valores en un intervalo (a, b), diremos que $X \in \mathrm{U}(a, b)$ si su función de densidad está dada por:

$$f(x) = \frac{1}{a-b} \; si \; x \in (a, b)$$

$$f(x) = 0 \qquad en \; cualquier \; otro \; caso$$

La función de densidad de probabilidad está dada por:

$$\int_a^b f(x)dx = 1$$

Su función de distribución es:

$$F(X) = 0 \;\; si \;\; x < a$$

$$F(x) = \frac{x-a}{b-a} \qquad si \; a \leq x \leq b$$

$$F(x) = 1 \; si \;\; x > b$$

El valor esperado es la media:

$$\mu = E(x) = \frac{b+a}{2}$$

La varianza está dada por la siguiente fórmula:

$$Var(x) = \frac{(b-a)^2}{12}$$

3.4.3 Distribución de probabilidad exponencial

Una variable aleatoria X tiene una distribución exponencial y se denomina variable aleatoria exponencial si y sólo si su función de densidad de probabilidad está dada por:

$$f(x) = \lambda e - \lambda x \quad si\ x > 0$$

Donde:

$$\lambda = \frac{\mathbf{1}}{\boldsymbol{E(X)}}$$

$$f(0) = 0 \quad en\ cualquier\ otro\ caso$$

Se considera una función de densidad de probabilidad cuando cumple con lo siguiente:

$$f(x) >= 0$$

Lo cumple por ser una función exponencial.

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Demostración

$$\int_{0}^{\infty} \lambda e^{\{-\lambda x\}}\, dx$$

$$\lim_{b \to \infty} \int_{0}^{b} \lambda e^{\{-\lambda x\}}\, dx$$

Integral que se resuelve por cambio de variable.

Sea:

$$\boldsymbol{t = e^{-\lambda x}}$$

$$\boldsymbol{\frac{dt}{dx} = -\lambda e^{-\lambda x}}$$

$$\boldsymbol{dt = -\lambda e^{-\lambda x} dx}$$

Entonces:

$$\lim_{b\to\infty}\int_0^b e^{-\lambda x}(\lambda)dx = -\lim_{b\to\infty}\int_0^b e^{-\lambda x}(-\lambda)dx$$

Realizando el cambio de variable:

$$-\lim_{b\to\infty}\int_0^b e^{t}dt = -\lim_{b\to\infty}(e^{t})|_0^b$$

Sustituyendo:

$$\boldsymbol{t = e^{-\lambda x}}$$

$$-\lim_{b\to\infty}(e^{-\lambda x})\Big|_0^b = -\lim_{b\to\infty}\left(\frac{1}{e^{\lambda x}}\right)\Big|_0^b$$

Evaluando la integral:

$$-\left[\lim_{b\to\infty}\left(\frac{1}{e^{\lambda b}}\right) - \lim_{b\to\infty}\left(\frac{1}{e^{\lambda}}\right)\right] = -\left[0 - \lim_{b\to\infty}\left(\frac{1}{e^{0}}\right)\right]$$

$$-\left[0 - \lim_{b\to\infty}\frac{1}{1}\right]$$

$$= \lim_{b\to\infty}(1)$$

$$= 1$$

Con lo cual queda demostrado la función de densidad de probabilidad.

3.4.3.1 Función de distribución

Demostración

$$F(x) = \int_0^x \lambda e^{-\lambda t}dt$$

La integral se resuelve por cambio de variable.

Sea:

$$\boldsymbol{w = e^{-\lambda t}}$$

$$\boldsymbol{\frac{dw}{dx} = -\lambda e^{-\lambda t}}$$

$$\boldsymbol{dw = -\lambda e^{-\lambda t} dt}$$

Entonces:

$$F(x) = \int_0^x \lambda e^{-\lambda t} dt$$

$$F(x) = \int_0^x e^{-\lambda t} (\lambda)\, dt$$

$$F(x) = -\int_0^x e^{-\lambda t} (-\lambda)\, dt$$

Realizando el cambio de variable:

$$F(x) = -\int_0^x e^{-\lambda t} (-\lambda)\, dt$$

$$F(x) = -\int_0^x e^{w}\, dw$$

$$F(x) = -e^{w}\Big|_0^x$$

Evaluando la integral:

$$F(x) = -e^{w}\Big|_0^x$$

$$F(x) = -e^{-\lambda t}\Big|_0^x$$

$$F(x) = -\left[e^{-\lambda x} - e^{-\lambda(0)}\right]$$

$$F(x) = -\left[e^{-\lambda x} - e^{0}\right]$$

$$F(x) = 1 - e^{-\lambda x}$$

Por lo tanto, la función de distribución es:

$$F(x) = 1 - e - \lambda x \ldots si\ x \geq 0$$

$$F(x) = 0 \quad si\ x < 0$$

La media y la varianza de la distribución exponencial son, respectivamente:

$$E(x) = \frac{1}{\lambda}$$

$$Var(x) = \frac{1}{\lambda^2}$$

3.4.4 Distribución de probabilidad normal

El 12 de noviembre de 1733, Abraham DeMoivre desarrolló la ecuación matemática de la curva normal, la cual muestra la distribución de los datos de forma simétrica alrededor de la media (μ). De igual manera proporcionó una base sobre la cual se fundamenta una gran parte de la teoría de la estadística inductiva. A la distribución normal se le llama también Distribución Gaussiana, en honor a Karl Friedrich Gauss.

Por tanto, una variable aleatoria X tiene una distribución normal con parámetros μ y σ^2, siendo μ un número real cualquiera y $\sigma > 0$, siendo su función de densidad de probabilidad la siguiente:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Con:

$$-\infty < x < \infty$$

$$-\infty < \mu < \infty$$

$$\sigma > 0$$

Para ser una función de densidad de probabilidad, debe de satisfacer las siguientes condiciones:

1. Que $f(x) \geq 0\ ;\ \forall x \in \mathbb{R}$.

   Que por ser una función exponencial lo cumple.

2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$

Propiedades

1. Es unimodal.

2. La media, la mediana y la moda poseen el mismo valor.

3. El dominio de *f*(*x*) son todos los números reales y su imagen está contenida en los números reales positivos.

4. Es simétrica respecto de la recta x - μ

   Esto se debe a que $f\ (\mu\ +\ x)\ =\ f\ (\mu - x)$

5. Tiene una asíntota horizontal en y = 0

   En efecto y = 0 es una asíntota horizontal, ya que $\lim_{x\to\infty} f(x) = 0$

6. Alcanza un máximo absoluto en el punto:

$$\left(\mu, \frac{1}{\sqrt{2\pi}\sigma}\right)$$

Demostración

Considere:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Derivando con respecto a μ

$$f'(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left(-\frac{2(x-\mu)}{2\sigma^2}(-1)\right)$$

$$f'(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left(\frac{2(x-\mu)}{\sigma^2}\right)$$

Igualando a cero:

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left(\frac{2(x-\mu)}{\sigma^2}\right) = 0$$

De donde resulta:

$$x - \mu = 0$$

$$x = \mu$$

Obteniendo la segunda derivada:

$$f'(x)(x)\left(\frac{x-\mu}{\sigma^2}\right)\left(\frac{x-\mu}{\sigma^2}\right)(x)\left(-\frac{1}{\sigma^2}\right)$$

$$f'(x)(x)\left(\frac{(x-\mu)^2}{\sigma^4}\right)(x)\left(-\frac{1}{\sigma^2}\right)$$

$$f'(x)(x)\left[\left(\frac{(x-\mu)^2}{\sigma^4}\right)-\frac{1}{\sigma^2}\right]$$

$$f'(x)(x)\left(\frac{1}{\sigma^2}\right)\left[\left(\frac{(x-\mu)^2}{\sigma^2}\right)-1\right]$$

De donde resulta:

$$f'(x)\frac{1}{\sigma^2}$$

Entonces:

$$f\ (\mu) < 0$$

Y como:

$$x = \mu$$

Por lo tanto:

$$f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$$

Lo que prueba que el máximo se encuentra en:

$$\left(\mu, \frac{1}{\sqrt{2\mu}\sigma}\right)$$

7. Es creciente en el intervalo (-∞, μ) y decreciente en (μ < ∞):

   Si $x < \mu$, es $f'(x) > 0$, entonces la función es creciente.

   Si $x > \mu$, es f´(x) <= 0, entonces la función es decreciente.

8. Posee dos puntos de inflexión en:

$$x = \mu - \sigma$$

$$x = \mu + \sigma$$

   Considerando la segunda derivada, la obtenida en el punto seis, e igualando a cero:

$$f(x)\left(\frac{1}{\sigma^2}\right)\left[\frac{(x-\mu)^2}{\sigma^2} - 1\right] = 0$$

   Despejando:

$$\frac{(x-\mu)^2}{\sigma^2} = 1$$

   De donde resulta:

$$(x - \mu)^2 = \sigma^2$$

$$x - \mu = \sigma \; ; \; x - \mu = -\sigma$$

   Por lo tanto:

$$x = \sigma + \mu \; ; \; x = \mu - \sigma$$

9. Los parámetros de μ y σ^2 son la media y la varianza respectivamente.

3.4.5 Distribución de probabilidad normal tipificada

Si X tiene una distribución normal con la media y la desviación estándar, entonces:

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

Donde:

X es la variable de interés.

μ es la media.

σ es la desviación estándar.

z es el número de desviaciones estándar de X respecto a la media de esta distribución.

La distribución normal estándar tiene un valor de media igual a cero y una varianza igual a 1, y se denota como N~(0,1).

Propiedades

1. Su dominio son todos los números reales y su imagen son los números reales positivos.

2. Es simétrica respecto al eje de ordenadas.

3. Tiene una asíntota horizontal en y = 0.

4. Alcanza un máximo absoluto en el punto $(0, \frac{1}{\sqrt{2\pi}})$

5. Es creciente en el intervalo (-∞ , 0) y decreciente en el intervalo (0, ∞).

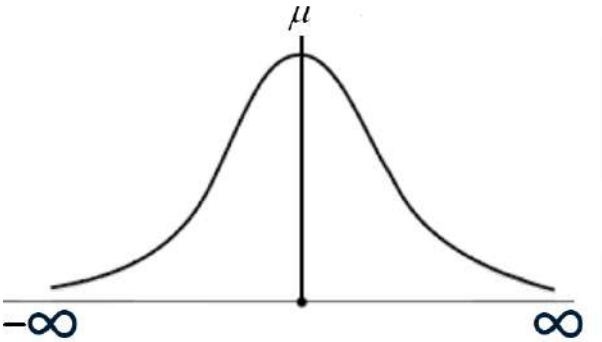


Figura 3.4 Distribución normal.
Fuente: elaboración propia a partir del procesamiento de datos.

6. Posee dos puntos de inflexión en x = -1 y en x = 1, respectivamente.

Áreas de la normal

Finalmente, podemos resumir la curva de una distribución normal de datos con las siguientes afirmaciones:

1. Aproximadamente el 68% de todos los valores se encuentran dentro de una desviación estándar.
2. Aproximadamente el 95.5% de todos los valores se encuentran dentro de dos desviaciones estándar.
3. Aproximadamente el 99.7% de todos los valores se encuentran dentro de tres desviaciones estándar.

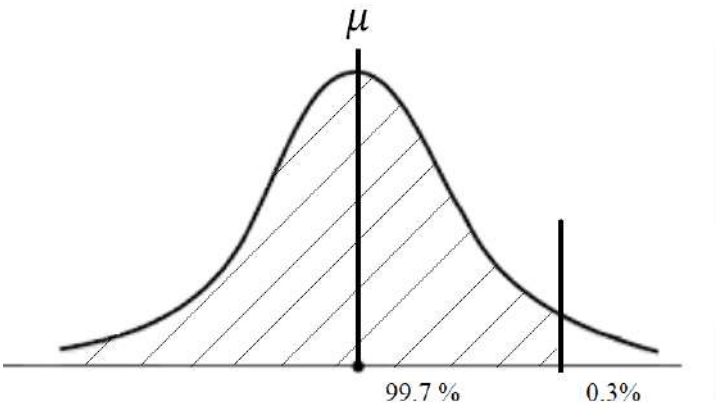


Figura 3.5 Áreas de la normal.
Fuente: elaboración propia a partir del procesamiento de datos.

Los siguientes dos videos (14 y 15) muestran un ejercicio que se resolverá a partir de la función de distribución Binomial y Normal.


Recursos Didácticos Digitales (Manual y Videos) para
Estadística, Probabilidad e Introducción a la Econometría
UNAM-DGAPA-PAPIME PE306125
Video 15
Función de distribución normal
Autor: Edmar Ariel Lezama Rodríguez
https://youtu.be/6sIYCM7hxRc
D.R.©, 2025, UNAM. CC BY-NC-ND

CAPÍTULO 4

Pruebas de hipótesis

En estadística y probabilidad, una prueba de hipótesis sirve para verificar la veracidad de un parámetro o dato en particular, así como para contrastar datos o cifras entre sí.

Para Quevedo (2011), una prueba de hipótesis se define como el establecer si el azar tiene que ver o no con la diferencia de dos datos, es decir, contrastar un dato de la muestra contra un dato de la población total.

Otro autor como Frías-Navarro (2025), plantea que una prueba de hipótesis debe partir de conocer datos sobre dos poblaciones, los cuales se pueden obtener de la población total y de una parte de esa población.

En el caso de Mendenhall (2010), una prueba de hipótesis puede ayudarnos a establecer diferencias entre dos grupos poblacionales a partir de sus parámetros principales, tales como la media.

Por tanto, para este texto se plantea que una prueba de hipótesis es la metodología que nos permitirá conocer si un parámetro corresponde a un determinado grupo poblacional (distinguir entre población total y muestra), contrastar dos datos para saber si tienen el mismo comportamiento, si provienen del mismo grupo poblacional o comprueban directamente cierto comportamiento.

Dentro de la teoría relacionada a las pruebas de hipótesis existen dos conceptos clave que son la hipótesis nula y la hipótesis alternativa.

La hipótesis nula H_0 es la que se somete a prueba y sobre ella se hace la decisión; para los propósitos de la prueba se asume como verdadera, mientras que la hipótesis alternativa H_a es la hipótesis planteada como opuesta a la prueba H_0 (Fallas, 2012).

Para ejemplificar el concepto de hipótesis nula e hipótesis alternativa se plantea lo siguiente:

Ejemplos

Supongamos que la persona docente de alguna asignatura en la Facultad de Economía dice que no habrá reprobados en su curso (H_0) , mientras que la segunda persona docente que se hace cargo del grupo dice que por lo menos habrá 5 reprobados (H_a).

Se plantean las hipótesis.

$$H_0 = \mu = 0 \qquad | \qquad H_a = \mu \geq 5$$

En otro ejemplo, supongamos que en una fábrica de focos se sabe que cada bombilla tiene una vida útil de 1,000 horas, mientras que en la empresa competidora los focos duran en promedio 980 horas, por lo que si el objetivo es conocer de donde proviene una bombilla a partir de la vida útil, se plantea:

$$H_0 = \mu = 1000$$

$$H_a = \mu = 980$$

Por tanto, podemos decir que la hipótesis nula hace referencia al dato principal, a la variable que se quiere probar o a la afirmación de nuestro problema.

Una vez hecha la distinción entre hipótesis nula e hipótesis alternativa es necesario plantear la definición formal, tal como se muestra a continuación.

Sea X= ($X_1, X_2, \ldots, X_n$) una muestra aleatoria proveniente de una población cuya distribución está denotada por:

$$\theta \in \ominus \subseteq \mathbb{R}^k$$

la cual define el espacio paramétrico de la hipótesis; las hipótesis planteadas para dicha muestra son:

- Hipótesis nula $H_0: \theta \in \ominus_0$
- Hipótesis alternativa $H_1: \theta \in \ominus_1$

Donde:

$$\ominus_0 \cap \ominus_1 = \emptyset \qquad\qquad \ominus \cup \ominus_1 = \ominus$$

Con una regla de decisión que está dada por la función:

$$\varphi: X \to 0,1$$

Donde:

- $\varphi(x) = 1: \ se \ rechaza \ H_0$
- $\varphi(x) = 0: \ se \ rechaza \ H_1$

Es probable que en este punto el lector tenga la inquietud de si es posible equivocarse en la decisión sobre una hipótesis, es decir, no rechazar la hipótesis nula cuando es falsa o rechazarla cuando es verdadera. Estas situaciones corresponden a los errores tipo I y tipo II. El error tipo I ocurre cuando se rechaza una hipótesis nula que es verdadera, mientras que el error tipo II ocurre cuando no se rechaza una hipótesis nula que es falsa. Formalmente, el error tipo I se comete cuando se rechaza la hipótesis nula H_0 siendo verdadera, y su probabilidad se define como:

$$\alpha = \mathbb{P}(H_0 \,|H_0 \ es \ verdadera)$$

α es conocido como el nivel de significancia, mientras que el error tipo II se genera cuando no se rechaza H_0 cuando ésta es falsa, es decir:

$$\alpha = \mathbb{P}(H_0|H_0 \ es \ falsa)$$

En la tabla 4.1 observamos de manera mejor, cuando se ocurren dichos errores.

Tabla 4.1 Resultados en la elección de una hipótesis. Fuente: elaboración propia.

	H_0 verdadera	H_0 falsa
Rechazar H_0	Error tipo I	Elección correcta
Rechazar H_1	Elección correcta	Error tipo II

En este punto sabemos diferenciar entre hipótesis nula e hipótesis alternativa, pero es necesario introducir un nuevo concepto al tema, el cual está asociado al nivel de significancia.

El nivel de significancia es una regla que permite afirmar que la diferencia observada entre dos o más grupos de datos (también conocidos como sets) es el resultado del efecto del "tratamiento" y no del azar. Con frecuencia se declaran como significativas aquellas diferencias con una probabilidad inferior a $0{,}05$ (es decir 5%) de observarse en forma aleatoria.

En algunos textos de estadística se recomienda utilizar un asterisco (*) para designar diferencias significativas a un 5% ($P<0.05$), dos asteriscos (**) para designar diferencias significativas al 1% ($P<0.01$) y tres asteriscos (***) para designar diferencias significativas al 0.1% ($P<0.001$)(Fallas, 2012).

Dentro de las pruebas de hipótesis existen tres tipos: la de dos colas, la de cola izquierda y la de cola derecha; cada caso se muestra a continuación.

4.1 Prueba de hipótesis

DE DOS COLAS

La prueba (de hipótesis) de dos colas se utiliza cuando se quiere detectar cualquier cambio significativo respecto al valor de la H_0, es decir:

$$H_0 = \mu = \mu_0 \qquad\qquad H_a = \mu \neq \mu_0$$

Este tipo de prueba se denomina también bilateral, ya que la H_0 se rechazará si el valor del estadístico de "p" prueba que se ubica en cualquiera de las zonas de rechazo de H_0, tal como se observa en la figura 4.1.

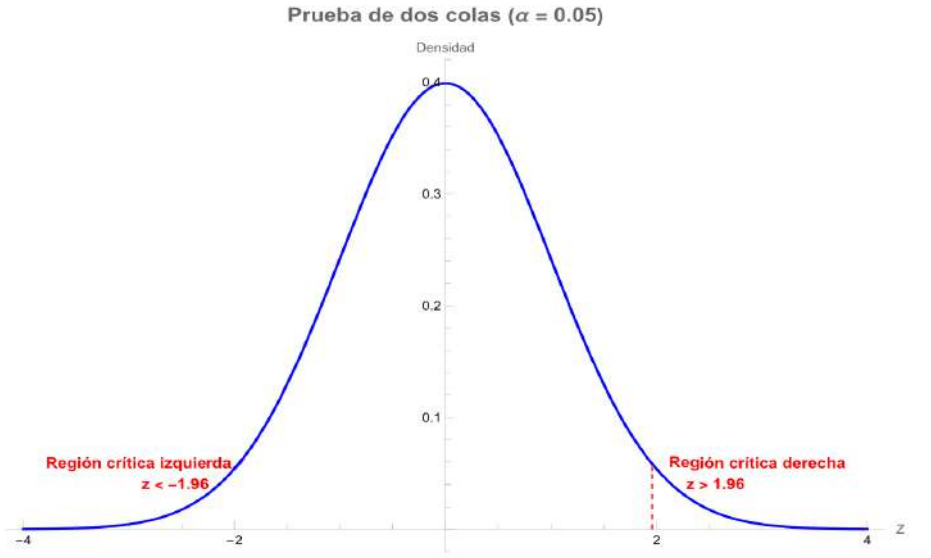


Figura 4.1 Pruebas de dos colas. Fuente: elaboración propia con software Mathematica.

La zona de rechazo es aquel conjunto de valores para los que rechazamos a la hipótesis nula (H_0) y dicha región es conocida también como región crítica.

Para calcular, si es que se rechaza o no se rechaza la H_0, usaremos la siguiente ecuación:

$$Z = \frac{X_M - \mu_0}{\sigma / \sqrt{n}}$$

Por lo que podemos plantar el siguiente intervalo: si $Z \in (-1.92, 1.92)$, no se rechaza H_0, por lo tanto[1]:

$$-1.92 < Z < 1.92$$

4.2 Prueba de hipótesis

DE COLA IZQUIERDA

Una prueba (de hipótesis) de cola izquierda, también llamada prueba de cola inferior es un tipo de prueba de hipótesis estadística donde la región de rechazo se encuentra en la cola izquierda de la distribución de probabilidad.

Se utiliza cuando la hipótesis alternativa establece que el parámetro poblacional es menor que el valor supuesto en la hipótesis nula.

Sea $X = (X_1, X_2, \dots, X_n)$ una muestra aleatoria simple de tamaño n con distribución tal que $f(x;\theta)$, donde $\theta \in \ominus \in \mathbb{R}$ es el parámetro de interés. Plantearemos la siguiente hipótesis para nuestro experimento de una cola:

$$\begin{cases} H_0: \theta \geq \theta_0 \\ H_1: \theta < \theta_0 \end{cases}$$

Una prueba de hipótesis unilateral de cola izquierda con nivel de significancia $\alpha \in (0,1)$ es una función de decisión $\varphi: \mathbb{R}^n \to \{0,1\}$ tal que:

$$\begin{cases} \varphi(x) = 1 & si\ se\ rechaza H_0 \\ \varphi(x) = 0 & si\ no\ se\ rechaza\ H_0 \end{cases}$$

Con la siguiente regla de decisión, definida por la zona crítica $C_\alpha \subseteq \mathbb{R}^n$, tal que:

$$\varphi = \begin{cases} 1 & si\ x \in C_\alpha \\ 0 & si\ x \notin C_\alpha \end{cases}$$

[1] Normalmente se usa α = 0.05

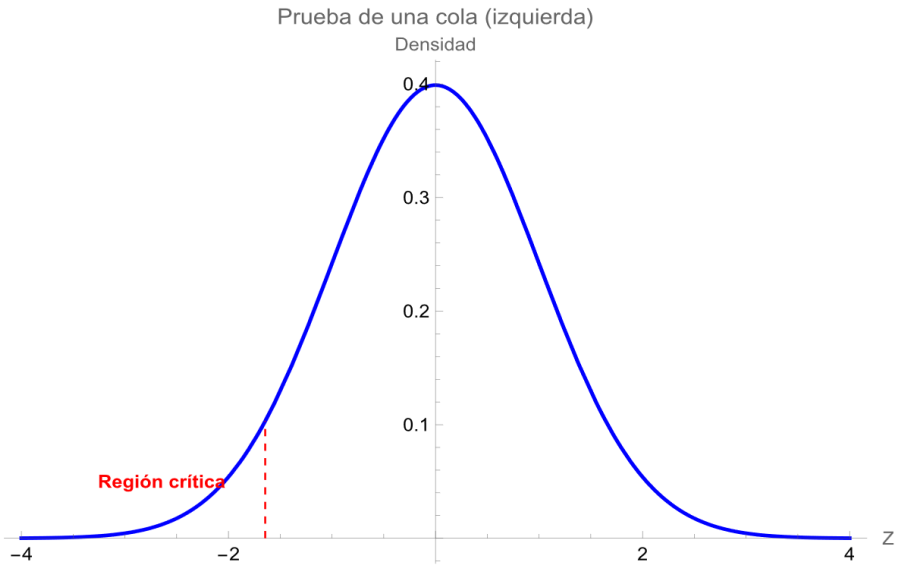


Figura 4.2 Representación gráfica de la prueba de cola izquierda. Fuente: elaboración propia con software Mathematica.

4.3 Prueba de hipótesis

DE COLA DERECHA

Una prueba (de hipótesis) de cola derecha es un contraste estadístico en el que la hipótesis alternativa tiene la forma:

$$H_1 : \mu > \mu_0$$

Se utiliza para determinar si la media (u otro parámetro) de una población es mayor que un valor específico. La región crítica (o de rechazo) se encuentra en la cola derecha de la distribución del estadístico de prueba, y la decisión de rechazar la hipótesis nula se toma si el estadístico cae por encima del valor crítico correspondiente al nivel de significancia α. Una prueba de hipótesis de cola derecha es un procedimiento estadístico diseñado para contrastar una hipótesis nula H_0, que típicamente establece que un parámetro poblacional θ (por ejemplo, la media μ) es igual a un valor específico θ, frente a una hipótesis alternativa H_1 que postula que dicho parámetro es estrictamente mayor, es decir, $H_1 : \theta > \theta_0$

El procedimiento consiste en calcular un estadístico de prueba T cuya distribución bajo H_0 es conocida, y determinar la probabilidad de obtener un valor tan extremo o más extremo que el observado, bajo la suposición de que H_0 es verdadera.

Esta probabilidad se denomina *valor-p.* La región de rechazo se ubica en la cola derecha de la distribución de T, y H_0 se rechaza si el estadístico de prueba cae dentro de dicha región, es decir, si $T > t_\alpha$, donde t_α es el cuantil de orden 1 - α de la distribución bajo H_0, siendo α el nivel de significancia previamente fijado.

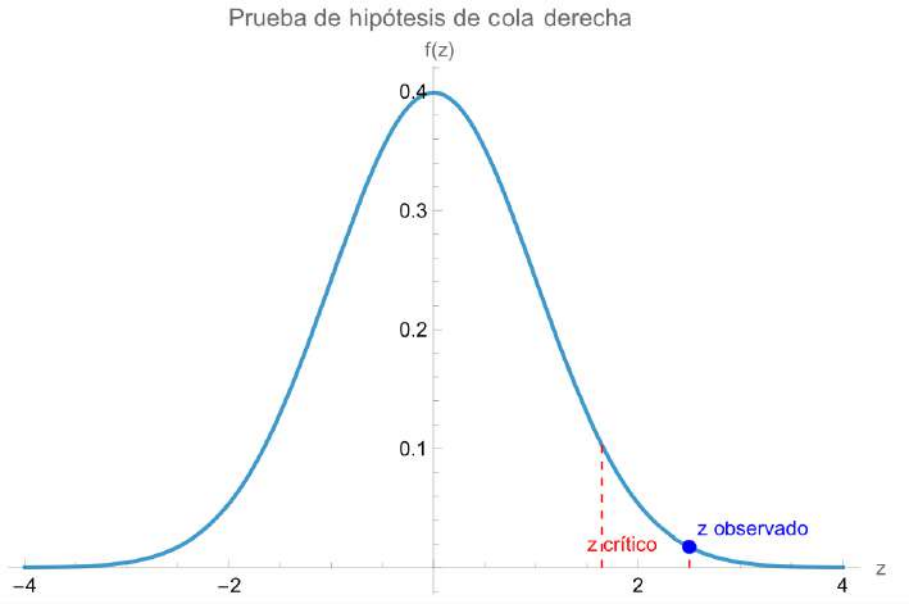


Figura 4.3 Representación gráfica de la prueba de cola derecha. Fuente: elaboración propia con software Mathematica.

El video 16 muestra una explicación de las distintas pruebas de hipótesis, así como la diferencia entre una hipótesis nula y alternativa; también se muestra cómo plantear un ejercicio a partir de la base de datos entre el salario percibido entre hombres y mujeres. Se sugiere replicar ese ejercicio para cualquier año de la base de datos.

https://youtu.be/07U0KpjiykY

El video 17 muestra una prueba de hipótesis llamada Test de Chow la cual sirve para plantear un cambio estructural entre poblaciones. Se sugiere usar la base de datos y calcular el cambio estructural entre las horas trabajadas entre mujeres y hombres para cualquier año de los datos proporcionados.

https://youtu.be/Yn8w8Zmf6D4

5

CAPÍTULO

Introducción a la econometría, a través de los mínimos cuadrados ordinarios


$y = b_0 + b_1 x$

Antes de empezar a definir conceptos y términos clave de la econometría, es necesario construir una definición propia, la cual servirá de guía para este texto.

Para Wooldridge (2010), la econometría se basa en el desarrollo de métodos estadísticos que se utilizan para estimar relaciones económicas, probar teorías económicas y evaluar e implementar políticas públicas y de negocios.

En el caso de Baltagi (2024), la econometría es una mezcla entre matemáticas, estadística y teoría económica para poder hacer predicciones sobre variables económicas que afectan a un país o a un grupo de personas en específico.

Autores como Maitra (2025) definen a la econometría como la integración entre la teoría económica, la modelación matemática y la inferencia estadística para analizar datos económicos del mundo real.

Otra definición clásica sobre econometría es la que ofrece Spanos (1986) al mencionar que la econometría tiene como objetivo dar contenido empírico a las relaciones económicas para probar teorías económicas, hacer pronósticos, tomar decisiones y para la evaluación ex post de decisiones y políticas.

Para Mignon (2022) la econometría es una disciplina con un fuerte componente operativo que permite cuantificar un fenómeno, establecer una relación entre diversas variables, validar o invalidar una teoría y evaluar los fenómenos de una política pública.

Otro autor clásico en temas de econometría es Gujarati (2021), quien la define como la combinación de la economía, la estadística inferencial y las matemáticas para analizar un fenómeno económico.

Por tanto, para fines de este texto definiremos a la econometría como el uso de técnicas probabilísticas y estadísticas para evaluar ecuaciones que serán estimadas y nos otorgarán coeficientes que medirán las relaciones entre la variable dependiente con las variables independientes.

Cabe mencionar que en la econometría es posible realizar un análisis a través de los datos sin importar si se trata de una serie temporal o de un momento en el tiempo.

Si los datos a emplear hacen referencia a un periodo de tiempo, se dice que la técnica econométrica a utilizar es una serie de tiempo o modelo panel; si los datos solo reflejan un problema en un momento específico en el tiempo (por ejemplo, un mes o trimestre en particular, el promedio de un año, entre otros) se dice que se trata de una sección cruzada o de corte transversal.

Para fines de este capítulo, utilizaremos datos de sección cruzada a través de la metodología de mínimos cuadrados ordinarios (MCO) (ver video 18).

https://youtu.be/jpqckDYVOzI

5.1 REPASO de álgebra matricial

Dado que el planteamiento de un problema de MCO requiere un planteamiento matricial, consideramos importante hacer un breve repaso sobre el tema.

El álgebra matricial es una subrama del álgebra lineal que estudia y analiza las propiedades y operaciones realizadas sobre las matrices.

Una matriz, por definición, es un arreglo de números cualesquiera organizados por n columnas y m filas; normalmente se utilizan para analizar un sistema de ecuaciones, así como para transformaciones lineales y de datos estadísticos.

5.1.1 Definición formal de una matriz

Una matriz $\boldsymbol{A}$ de dimensión $m \times n$ es una función, tal que

$$A: \{1,2,\dots,m\} \times \{1,2,\dots,n) \to \mathbb{F}$$

Donde $\mathbb{F}$ es un campo cualquiera y es representada de la siguiente manera.

$$A = \begin{bmatrix} a_{11} & b_{12} \\ a_{21} & b_{22} \end{bmatrix}$$

Recuerde que una forma abreviada para describir una matriz es mediante la expresión $A = [a_{ij}]$ donde $i = 1, 2,\dots, m$ y $j = 1, 2,\dots, n$, lo cual nos indica el orden o dimensión de la matriz ($m \times n$). Una matriz es cuadrada cuando $m = n$ y es rectangular cuando $m \neq n$.

5.1.2 Suma y resta de matrices

La condición obligatoria que se debe cumplir a la hora de sumar y restas matrices es que ambas deben tener el mismo número de columnas y filas.

Supongamos que $A = 3 \times 2$ y $B = 3 \times 3$; estas dos matrices no se pueden sumar ni restar. Esto es así ya que, tanto para la suma como para la resta se aplica el operador sobre los términos que ocupan el mismo lugar en las matrices y en este ejemplo el número de columnas es distinto.

Ejemplo

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; A + B = \begin{bmatrix} 1+1 & 2+2 \\ 3+3 & 4+4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; B = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}; A - B = \begin{bmatrix} 1-4 & 2-3 \\ 3-2 & 4-1 \end{bmatrix} = \begin{bmatrix} -3 & -1 \\ 1 & 3 \end{bmatrix}$$

5.1.3 Matriz transpuesta

Transponer una matriz es cambiar las filas por las columnas, y viceversa. Se debe respetar el orden de presentación de los elementos, es decir la primera columna de la matriz A corresponde a la primera fila de la matriz transpuesta y así de forma subsecuente.

Ejemplo

$$A = \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix}$$

$$A^t = \begin{pmatrix} a & c & e \\ b & d & f \end{pmatrix}$$

5.1.4 Multiplicación de matrices

Para poder multiplicar dos matrices, la primera debe tener el mismo número de columnas que filas tiene la segunda. La matriz resultante del producto quedará con el mismo número de filas de la primera y con el mismo número de columnas de la segunda.

Ejemplo

$$(5 \times 3) \times (3 \times 2) = (5 \times 2)$$

Supongamos que $A=(a_{ij})$ y $B=(b_{ij})$ son matrices tales que el número de columnas de A coincide con el número de filas de B; es decir, A es una matriz $m \times p$ y B una matriz $p \times m$. Entonces, el producto AB es la matriz $m \times n$ cuya entrada ij se obtiene multiplicando la fila de i por la columna j de B.

$$\begin{pmatrix} \beta & \alpha \\ \sigma & \theta \end{pmatrix} \times \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} = \begin{pmatrix} \beta a_1 + \alpha b_1 & \beta a_2 + \sigma b_2 & \beta a_3 + \sigma b_3 \\ \sigma a_1 + \theta b_1 & \sigma a_2 + \theta b_2 & \sigma a_3 + \theta b_3 \end{pmatrix}$$

El procedimiento puede llegar a ser algo confuso, pero en esencia es completamente sencillo; en siguiente ejemplo lo explicaremos de manera numérica.

Ejemplo

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \times \begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1*4 + 2*2 & 1*3 + 2*1 \\ 3*4 + 4*2 & 3*3 + 4*1 \end{pmatrix} = \begin{pmatrix} 8 & 5 \\ 20 & 13 \end{pmatrix}$$

Al igual que en álgebra el producto de matrices es asociativo, por lo tanto decimos que $A \in M_{n\times m}(F), B \in M_{m\times r}(F), C \in M_{r\times s}(F)$, por lo tanto:

$$(AB)C = A(BC) \in M_{n\times s}(F)$$

5.1.5 Escalar por una matriz

El producto de una matriz por un escalar es en parte una operación sencilla pero no por eso, la pasaremos por alto. Sea K un escalar perteneciente a cualquier campo denotado por $\mathbb{F}$ por lo que decimos que $k \in \mathbb{F}$, entonces decimos que un escalar por una matriz se denota de la siguiente manera:

$$k \times A$$

Ejemplo

$$3A = 3 \times \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 3*1 & 3*2 \\ 3*3 & 3*4 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 9 & 12 \end{pmatrix}$$

5.1.6 Matriz inversa

Decimos que una matriz A tiene inversa, si $\exists B$ tal que $AB = I_n = BA$; la matriz debe cumplir con el requisito de ser una matriz cuadrada de orden $m \times m$, donde $m \neq 0$.

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^{-1} = \begin{pmatrix} cos(-\theta) & -sin(-\theta) \\ sin(-\theta) & cos(-\theta) \end{pmatrix}$$

Por lo tanto

$$\begin{pmatrix} cos(\theta) & -sen(\theta) \\ sen(\theta) & cos(\theta) \end{pmatrix} \begin{pmatrix} cos(-\theta) & -sen(-\theta) \\ sen(-\theta) & cos(-\theta) \end{pmatrix} =$$

$$\begin{pmatrix} cos^2(\theta) + sen^2(\theta) & 0 \\ 0 & cos^2(\theta) + sen^2(\theta) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

5.1.6.1 Método Gauss-Jordan

El método de Gauss-Jordan se utiliza para transformar, de forma escalonada, un sistema de ecuaciones en otro equivalente y así obtener el valor de las incógnitas; en la matriz resultante se ponen los coeficientes de las variables y los términos independientes (separados por una recta).

Por ejemplo, se busca encontrar los valores para la matriz A, tal que:

$$A = \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}$$

Paso I

Usar el método Gauss-Jordán. Por lo tanto, obtenemos:

$$\left(\begin{array}{cc|cc} cos(\theta) & -sen(\theta) & 1 & 0 \\ sin(\theta) & cos(\theta) & 0 & 1 \end{array}\right)$$

Paso II

Dividimos la primera fila entre $cos(\theta)$

$$R_1 \leftarrow \frac{1}{cos(\theta)} R_1 = \left(\begin{array}{cc|cc} 1 & -tan(\theta) & sec(\theta) & 0 \\ sin(\theta) & cos(\theta) & 0 & 1 \end{array}\right)$$

Paso III

$$R_2 \leftarrow R_2 - sin(\theta) * R_1 = \left(\begin{array}{cc|cc} 1 & -tan(\theta) & sec(\theta) & 0 \\ 0 & \frac{1}{cos(\theta)} & -tan(\theta) & 1 \end{array}\right)$$

Paso IV

$$R_1 \leftarrow cos(\theta) * R_2 = \left(\begin{array}{cc|cc} 1 & -tan(\theta) & sec(\theta) & 0 \\ 0 & 1 & -sin(\theta) & cos(\theta) \end{array}\right)$$

Paso V

$$R\_1 \leftarrow R\_1 + tan(\theta) * R\_2 = \left(\begin{array}{cc|cc} 1 & 0 & cos(\theta) & sin(\theta) \\ 0 & 1 & -sin(\theta) & cos(\theta) \end{array}\right)$$

5.1.7 Método por determinantes

El método por determinantes es un método más amigable, en el sentido de que no es tan talachero como el anterior. Usaremos una sencilla ecuación, tal que.

$$A^{-1} = \frac{(A^d)^t}{|A|}$$

La única condición para que exista A^{-1} es que $|A|$ sea diferente de 0, donde $(A^{d})^{t}$ es la matriz adjunta que a su vez es la transpuesta.

Ejemplo

Sea:

$$A = \begin{pmatrix} 1 & -5 & 8 \\ 7 & 6 & 9 \\ -2 & 8 & 3 \end{pmatrix}$$

Calculamos el determinante, tal que:

$$|A| = a\begin{pmatrix} e & f \\ h & i \end{pmatrix} - b\begin{pmatrix} d & f \\ g & i \end{pmatrix} + c\begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

Tomando de ejemplo a nuestra matriz, la igualdad se expresaría:

$$A = \begin{pmatrix} 1 & -5 & 8 \\ 7 & 6 & 9 \\ -2 & 8 & 3 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

y la transpuesta queda como:

$$|A| = 1\begin{pmatrix} 6 & 9 \\ 8 & 3 \end{pmatrix} - (-5)\begin{pmatrix} 7 & 9 \\ -2 & 3 \end{pmatrix} + 8\begin{pmatrix} 7 & 6 \\ -2 & 3 \end{pmatrix}$$

Se asume que el lector sabe resolver determinantes, pero en caso de que no, existe una ecuación sencilla para resolverlo.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = |A| = ad - bc$$

Por lo tanto, decimos que:

$$|A| = 685$$

Para formar la nueva matriz, debemos usar determinantes 2×2; de nuestra matriz original, haremos las siguientes operaciones para generar una nueva matriz:

$$A_{11} = \begin{pmatrix} e & f \\ h & i \end{pmatrix}; \quad A_{12} = -\begin{pmatrix} d & f \\ g & i \end{pmatrix}; \quad A_{13} = \begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

$$A_{21} = -\begin{pmatrix} b & c \\ e & f \end{pmatrix}; \quad A_{22} = \begin{pmatrix} a & c \\ g & i \end{pmatrix}; \quad A_{23} = -\begin{pmatrix} a & b \\ g & h \end{pmatrix}$$

$$A_{31} = \begin{pmatrix} b & c \\ e & f \end{pmatrix}; \quad A_{32} = -\begin{pmatrix} a & d \\ c & f \end{pmatrix}; \quad A_{33} = \begin{pmatrix} a & b \\ d & e \end{pmatrix}$$

Calculando los determinantes, obtendremos una nueva matriz, la cual es llamada matriz de adjuntos; aún falta transponerla.

$$A^d = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{\{33]} \end{pmatrix} = (A^d)^t = \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}$$

Ya en este paso sólo es dividir por nuestro determinante de la matriz y obtener nuestra matriz inversa.

$$A^{-1} = \begin{pmatrix} -\frac{54}{685} & \frac{79}{685} & -\frac{93}{685} \\ -\frac{39}{685} & \frac{19}{685} & \frac{47}{685} \\ \frac{68}{685} & \frac{2}{685} & \frac{41}{685} \end{pmatrix}$$

5.2 TEOREMA de mínimos cuadrados ordinarios (MCO)

El método de mínimos cuadrados ordinarios es una técnica para ajustar la línea recta óptima a la muestra de las observaciones XY y comienza con:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon, \epsilon \sim N(0, \sigma^2), i = 1, \dots n$$

Los parámetros que minimizan la suma residual de cuadrados están dados por:

$$\beta_1 = \frac{\sum x_i y_i}{\sum x_i^2} \qquad \beta_0 = \bar{Y} + \beta_1 \bar{X}$$

Donde $\bar{Y}$ y $\bar{X}$ son las medias muestrales, $\sum x_i^2$ es la varianza muestral de x, $\sum x_i y_i$ son las covarianzas muestrales de x, y.

Demostración

La suma residual de cuadrados (SRC) está definida por:

$$SRC(\beta_0, \beta_1) = \sum_{i=1}^{n} \epsilon_i^2$$

Las derivadas de $SRC(\beta_0, \beta_1)$ con respecto a β_0 y β_1 son las siguientes:

$$\frac{dSRC(\beta_0,\beta_1)}{d\beta_0} = \sum_{i=1}^{n} 2(y_i - \beta_0 - \beta_1 x_i)(-1)$$

$$= -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)$$

$$\frac{dSRC(\beta_0,\beta_1)}{d\beta_1} = \sum_{i=1}^{n} 2(y_i - \beta_0 - \beta_1 x_i)(-x_i)$$

$$= -2\sum_{i=1}^{n}(x_i y_i - \beta_0 - \beta_1 x_i^2)$$

Se igualan a cero ambas derivadas:

$$0 = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)$$

$$0 = -2\sum_{i=1}^{n}(x_i y_i - \beta_0 - \beta_1 x_i^2)$$

Por lo tanto:

$$\widehat{\beta_1}\sum_{i=1}^{n} x_i + \widehat{\beta_0} n = \sum_{i=1}^{n} y_i$$

$$\widehat{\beta_1}\sum_{i=1}^{n} x_i^2 + \widehat{\beta_0}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i$$

A partir de estas ecuaciones, podemos derivar la estimación de la intersección, por lo tanto:

$$\widehat{\beta_0} = \frac{1}{n}\sum_{i=1}^{n} y_i - \widehat{\beta_1}\frac{1}{n}\sum_{i=1}^{n} x_i$$

$$= \bar{Y} - \beta_1 \hat{X}$$

De la segunda ecuación obtenemos:

$$\widehat{\beta_1} = \frac{\sum_{i=1}^{n} x_i y_i - \bar{y}\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 - \bar{x}\sum_{i=1}^{n} x_i}$$

El numerador se puede escribir como:

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

El denominador se puede escribir:

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

Ambos constituyen las estimaciones y la demostración de los parámetros del método de mínimos cuadrados ordinarios. Con esta demostración ya tenemos las fórmulas para calcular los estimadores β_0, y β_1 los cuales son:

$$\beta_0 = \bar{Y} + \beta_1 \bar{X} \qquad \beta_1 = \frac{\sum x_i y_i}{\sum x_i^2}$$

Donde $\sum x_i y_i$ es igual a $\sum(X_i - \bar{X})(Y_i - \bar{Y})$. Ahora daremos un ejemplo de cómo calcular y estimar una ecuación de mínimos cuadrados ordinarios.

Ejemplo

Cálculo de β_0 y β_1, a partir de la siguiente tabla:

	n	y	x	$y - Y$	$x - X$	$x\,y$	x^2
	1	40	6	17	12	204	144
	2	44	10	13	8	104	64
	3	46	12	11	6	66	36
	4	48	14	9	4	36	16
	5	52	16	5	2	10	4
	6	58	18	-1	0	0	0
	7	60	22	-3	-4	12	16
	8	68	24	-11	-6	66	36
	9	74	26	-17	-8	136	64
	10	80	32	-23	-14	322	196
$\sum$	n=10	570	180	0	0	956	576
Promedio				$\bar{Y} = 57$		$\bar{X} = 18$	

Tabla 5.1 Valores correspondientes al ejemplo presentado. Fuente: elaboración propia.

Recordemos que (y - $\bar{Y}$) es igual a y, (x - $\bar{X}$) es igual a x; dichos valores se calculan de la siguiente manera.

Tabla 5.2 Cálculo de $(y - \underline{Y})$ y $(x - \underline{X})$
Fuente: elaboración propia.

y	x	$(y - Y)$	$(x - X)$
40	6	17	12
44	10	13	8
46	12	11	6
48	14	9	4
52	16	5	2
58	18	-1	0
60	22	-3	-4
68	24	-11	-6
74	26	-17	-8
80	32	-23	-14
$\bar{Y} = 57$	$\bar{X} = 18$	0	0

Los valores 57 y 18 son los promedios de dichas variables; para calcular $(x - \bar{X})$ y $(y - \bar{Y})$ sólo hacemos la diferencia entre nuestro primer valor observado hasta n menos el promedio de nuestra variable observada. Ya solamente queda sustituir nuestros valores en las fórmulas previamente demostradas.

$$\beta_1 = \frac{956}{576} = 1.659722$$

$$\beta_0 = 57 + (1.65972)(18) = 86.8749$$

Por lo tanto:

$$\hat{Y} = 86.8749 + 1.6597X$$

Esta es la ecuación estimada de nuestro modelo correspondiente al ejemplo para calcular las pruebas de significancia de estimación de los parámetros (ver video 19).

5.3 REPRESENTACIÓN GRÁFICA de un modelo de MCO

Como ya se ha mencionado, un modelo de MCO busca establecer una relación entre variables independientes con una dependiente, lo cual se hace a partir de los coeficientes obtenidos. Gráficamente, la relación se establece a partir de diagramas de dispersión, es decir, se colocan las coordenadas en las cuales hay alguna relación entre dos variables y sobre esa nube de puntos se busca establecer una línea recta que ajuste de la mejor manera, tal como se muestra en la figura 5.1.

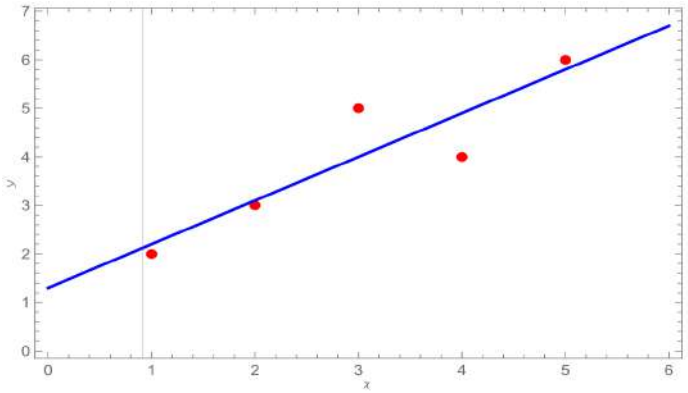


Figura 5.1 Representación gráfica de una regresión lineal. Fuente: elaboración propia con software Mathematica.

El eje Y siempre contendrá a la variable dependiente, mientras que el eje X albergará a la variable independiente.

Las relaciones a establecer pueden ser varias, por ejemplo, los años de escolaridad afectando al salario; la tasa de interés incentivando o no a la inversión; la tasa de inflación afectando al consumo, por tan solo citar algunos casos (ver video 20).

5.4 MÉTODO MATRICIAL del análisis de regresión

El análisis de regresión múltiple es un método eficiente para predecir fenómenos de índole económica, por ende, tiene varias formas de ser calculado; una de ellas es desde un enfoque matricial.

Iniciaremos con el análisis de regresión simple. Sea una ecuación de análisis de regresión simple tal que:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \qquad \forall i \in \mathbb{N} \; y \; i = 1 \; \ldots n$$

Por lo tanto, decimos que para i desde 1 hasta n:

$$Y_1 = \beta_0 + \beta_1 X_1 + \epsilon_2$$

$$Y_2 = \beta_0 + \beta_1 X_2 + \epsilon_2$$

Hasta:

$$Y_n = \beta_0 + \beta_1 X_n + \epsilon_n$$

Podemos formular la simple función de regresión lineal anterior en notación matricial, tal que:

$$Y = X\beta + \epsilon$$

Es decir, en lugar de escribir las n ecuaciones podemos utilizar la notación matricial mediante la cual nuestra función de regresión lineal se reduce a una declaración corta y simple. Para calcular la regresión lineal de manera simple, decimos que:

$$\beta = \begin{bmatrix} b_0 \\ b_1 \\ b_k \end{bmatrix} = (X'X)^{-1}X'Y$$

Donde X' es la matriz transpuesta, X es la matriz original, y $(X'X)^{-1}$ es la inversa del producto de las dos matrices ya mencionadas. La matriz X se denota de la siguiente manera:

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_n \end{pmatrix}$$

La matriz transpuesta de X se denota como:

$$X^t = \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_n \end{pmatrix}$$

Por lo tanto, decimos que:

$$X^tX = \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_n \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_1^2 \end{pmatrix}$$

Ejemplo

Supongamos la siguiente tabla:

Tabla 5.3
Datos del ejemplo.
Fuente: elaboración propia.

X	Y
1	2
2	3
3	5

Usando de referencia la fórmula establecida en párrafos anteriores $Y = X\beta + \epsilon$, decimos que:

$$Y = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}; \quad X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$$

Utilizando la ecuación siguiente decimos $\hat{\beta} = (X^tX)^{-1}X^tY$, por lo tanto,

$$X^tX = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$$

Luego calculamos X^tY:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 10 \\ 23 \end{bmatrix}$$

Ahora invertimos $(X^tX)^{-1}$; por lo tanto decimos que:

$$(X^tX)^{-1} = \frac{1}{6} \begin{bmatrix} 14 & -6 \\ -6 & 3 \end{bmatrix}$$

Para calcular $\hat{\beta}$:

$$\frac{1}{6}\begin{bmatrix}14 & -6\\ -6 & 3\end{bmatrix}\times\begin{bmatrix}10\\ 23\end{bmatrix}=\begin{bmatrix}\frac{1}{3}\\ \frac{3}{2}\end{bmatrix}$$

Y finalmente obtener:

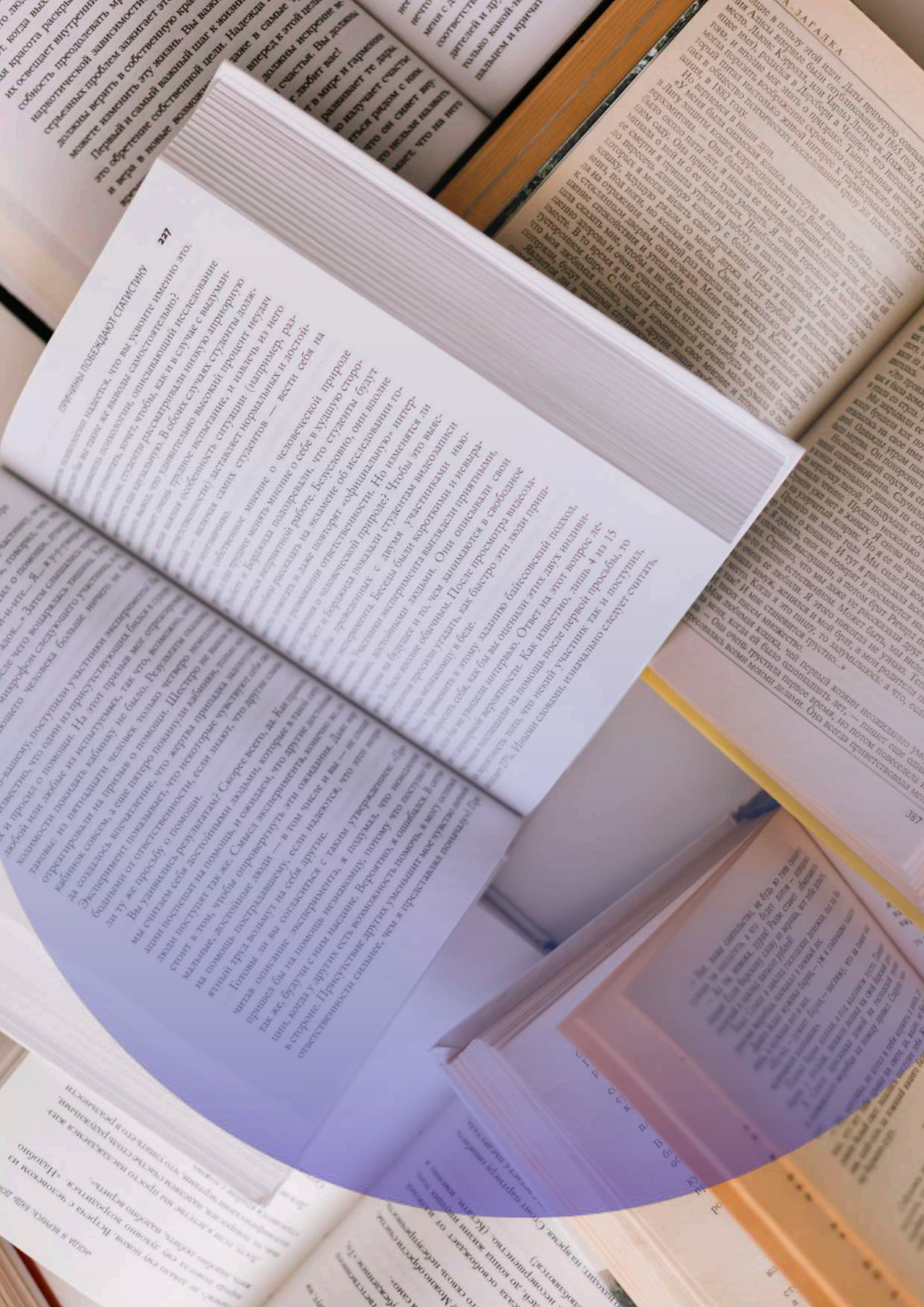
$$\hat{y}=\frac{1}{3}+\frac{3}{2}X$$

En los videos 21 a 24 se muestra el paso a paso de la construcción de un modelo de MCO relacionando con los conceptos teóricos de la economía. Como ejercicio se puede tener al salario como variable dependiente y a la escolaridad y las horas trabajadas como variables independientes y así conocer el comportamiento que tiene el salario ante las dos variables.

https://youtu.be/yAO386Qt88w

https://youtu.be/1CbwiaxLRcE

Referencias



1 | Apostol, T. (1991). Calculus. vol I. John Wiley & Sons.

2 | Apostol, T. (2019). Calculus II: cálculo con funciones de varias variables y álgebra lineal, con aplicaciones para ecuaciones diferenciales y probabilidad. Reverté.

3 | Baltagi, B. (Ed.). (2006). Panel data econometrics: Theoretical contributions and empirical applications. Emerald Group Publishing.

4 | Baltagi, B. (2008). Econometría. Springer Berlín Heidelberg. https://doi.org/10.1007/978-3-540-76516-5.

5 | Bismans, F., & Damette, O. (2025). Dynamic econometrics. Springer Books.

6 | Coll Serrano, V. (2023). Estadística descriptiva. Presentación en la Universitat de Valencia. https://www.uv.es/vcoll/Calculo_y_estadistica/Tema_2_estadistica.pdf.

7 | Fallas, J. (2012). Análisis de varianza. Comparando tres o más medias. https://www.academia.edu/39613853/AN%C3%81LISIS_DE_VARIANZA.

8 | Faraldo, P. (2012). Estadística y metodología de la investigación. Facultad de Enfermería. Universidad de Santiago de Compostela.

9 | Frías-Navarro, D. (2025). Redactar los resultados científicos de las hipótesis del informe de investigación (2024/2025). Universidad de Valencia.

10 | Geweke, J., Horowitz, J., & Pesaran, M. (2006). Econometrics: A bird's eye view. No. 1870. CESifo Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=947529.

11 | Huang, D. (1970). Introducción al uso de la matemática en el análisis económico. En el mundo del hombre. Economía y demografía. Siglo XXI.

12 | INEGI (2025). Glosario. https://www.inegi.org.mx/app/glosario/default.html?p=ENOE15.

13 | Intriligator, M. (1990). Modelos econométricos, técnicas y aplicaciones. Fondo de Cultura Económica.

14 | Kolmogorov, A. (1933). Grundbegriffe der Wahrscheinlichkeitsrechnung (en alemán). Berlin: Julius Springer.

15 | Laveaga, C. (2014). Álgebra superior. Curso completo. Universidad Nacional Autónoma de México. https://ulibros.com/algebra-superior-curso-completo-p6xrc.html.

16 | Maitra, S. (2025). A practical guide to static and dynamic econometric modelling. Contributions to Economics. https://link.springer.com/book/10.1007/978-3-031-86862-7.

17 | Maltsev, A. (1972). Fundamentos de álgebra lineal. Editorial Mir, Moscú, Rusia.

18 | Mendenhall, W. (2010). Introducción a la probabilidad y estadística. CENGAGE Learning.

19 | Mignon, V. (2024). Principles of econometrics: Theory and applications. Springer Nature. https://doi.org/10.1007/978-3-031-52535-3.

20 | Porras, A. (2017). Diplomado en análisis de información geoespacial. Conceptos básicos de estadística. Conacyt, Centro GEO.

21 | Quevedo, F. (2011). Medidas de tendencia central y dispersión. Estadística Aplicada a la Investigación en Salud. Medwave. Año XI, No. 3, Marzo 2011. Open Access, Creative Commons.

22 | Rendón-Macías, M., Villacís, M. y Miranda, M. (2016). Estadística descriptiva. https://revistaalergia.mx/ojs/index.php/ram/article/view/230.

23 | Rincón, L. (2007). Curso elemental de probabilidad y estadística. Universidad Nacional Autónoma de México.

24 | Rincón, L. (2014). Introducción a la probabilidad. pp. 217-232. Universidad Nacional Autónoma de México.

25 | Salvatore, D. & Reagle, D. (2002). Schaum's outline of theory and problems of statistics and econometrics. (2a ed.) McGraw-Hill.

26 | Vázquez-Alamilla, J. (2020). Inferencia estadística para estudiantes de ciencias. Prensas de Ciencias, UNAM.

27 | Wooldridge, J. (2010). Introducción a la econometría. Un enfoque moderno. CENGAGE Learning.

Ejercicios resueltos

A continuación, se presentan los ejercicios propuestos con el desarrollo de las respuestas considerando los temas abordados en el capítulo.

1. Se desea conocer la ubicación de una fábrica de muebles; la probabilidad de que dicha fábrica se ubique en el municipio A es del 70% , de que se ubique en el municipio B es del 40%, y de que se encuentre en ambos municipios es del 80%.

¿Cuál es la probabilidad de que se localice en ?

a.- Cualquiera de los dos municipios

b.- Ninguno de los municipios

Respuesta:

Primero, se debe separar cada una de las probabilidades para poder iniciar con los cálculos matemáticos: $P(A) = 0.70$, $P(B) = 0.40$, y $P(A \cap B) = 0.80$

a) Para calcular la probabilidad de que se localice en cualquiera de los dos municipios, decimos que

$$P(A) + P(B) - P(A \cap B)$$
$$0.70 + 0.40 - 0.80 = 0.30$$

b) Para determinar la probabilidad de que no se localice en alguno de los municipios es

$$1 - P(A \cap B)$$
$$1 - 0.80 = 0.20$$

2. En una compañía de 1500 empleados, 250 están en la sección de atención a clientes y 100 están en la sección de ventas al mayoreo. Calcular la probabilidad de que el empleado:

a) No se encuentren en la sección de atención a clientes.

b) No se encuentren en la sección de ventas al mayoreo.

c) Se encuentren en la sección de atención al cliente o en la sección de ventas al mayoreo.

Respuesta:

Se inicia planteando los datos. En este caso, para cada sector dividiremos el total de empleados de cada departamento entre el total de empleados, por tanto, decimos que

$$P(A) = \frac{250}{1500} = \frac{1}{6} = 0.16$$

$$P(B) = \frac{100}{1500} = \frac{1}{15} = 0.66$$

Ya con los datos individuales, se puede realizar el cálculo.

a) Probabilidad de que el empleado no se encuentre en la sección de atención a clientes.

$$1 - P(A)$$

$$1 - 0.16 = 0.84$$

En este caso, la probabilidad de que el empleado no se encuentre en la sección de atención a clientes, es de 0.84.

b) Probabilidad de que el empleado no se encuentre en la sección de ventas al mayoreo.

Al igual que en el ejemplo anterior, decimos que:

$$1 - P(B)$$

$$1 - 0.66 = 0.34$$

La probabilidad de que el empleado no se encuentre en la sección de atención a clientes, es de 0.34.

c) Probabilidad de que el empleado se encuentreestar en la sección de atención al cliente o en la sección de ventas al mayoreo.

Para calcular dicha probabilidad debemos sumar las primeras probabilidades calculadas anteriormente, por lo tanto, decimos que:

$$P(A \cup B) = P(A) + P(B) = \frac{250}{1500} + \frac{100}{1500} = \frac{7}{30} = 0.233$$

En este último caso, la probabilidad de que el empleado no se encuentre en la sección de atención a clientes, es de 0.233.

3. En una encuesta selectiva entre 180 personas que habían comprado un auto, 27 compraron un auto a gasolina y 72 uno eléctrico.

Calcular la probabilidad de elegir a una persona que haya comprado un auto eléctrico, uno a gasolina y ninguna de las opciones mencionadas.

Respuesta:

Para iniciar el cálculo, se obtiene cada probabilidad indicada; por lo tanto, decimos que:

$$P(Autos\ a\ gasolina) = \frac{27}{180} = 0.15$$

$$P(Autos\ eléctricos\ ) = \frac{72}{180} = 0.40$$

El cálculo de la probabilidad de que no se haya comprado ningún auto es de:

$$1 - [P(A) + P(B)] =$$

$$1 - \left[\frac{27}{180} + \frac{72}{180}\right] =$$

$$1 - \frac{99}{180} =$$

$$1 - 0.55 = 0.45$$

Por lo tanto, la probabilidad de que no se compre alguno de los otros coches es de 0.45.

4. La probabilidad de que un vuelo con programación regular despegue a tiempo es de 83%; la de que llegue a tiempo es 82% y la de que despegue y llegue a tiempo es del 78%. Calcular la probabilidad de que un vuelo elegido al azar.

Respuesta:

Al igual que en los ejercicios anteriores, se determinan las probabilidades que nos ofrece el enunciado del problema:

$$P(A) = 0.83$$
$$P(B) = 0.82$$
$$P(A \cap B) = 0.78$$

a) Probabilidad de que llegue a tiempo dado que despegó a tiempo.

$$P(B|A) = \frac{P(A \cap B)}{P(\mathrm{A})} = \frac{0.78}{0.83} = 0.93$$

En este caso, la probabilidad de que llegue a tiempo dado que despego a tiempo es de 0.93.

b) Probabilidad de que despegue a tiempo dado que llegó a tiempo.

$$P(A|B) = \frac{P(A \cap B)}{P(\mathrm{B})} = \frac{0.78}{0.82} = 0.95$$

Por lo tanto, la probabilidad de que despegue a tiempo dado que llegó a tiempo es de 0.95.

5. De un total de 200 adultos, se tiene la siguiente información:

Escolaridad	Hombre	Mujer
Primaria	38	45
Secundaria	28	50
Bachillerato	22	17

Si se selecciona aleatoriamente a una persona de ese grupo, calcular la probabilidad de:

a) Sea hombre, dado que tiene escolaridad de nivel secundaria.

b) Sea mujer, dado que tiene escolaridad de nivel bachillerato.

Respuesta:

Primero, se debe completar la tabla proporcionada en el ejercicio para posteriormente calcular las probabilidades solicitadas.

Escolaridad	Hombre	Mujer	Total
Primaria	38	45	83
Secundaria	28	50	78
Bachillerato	22	17	39
Total	88	112	200

Calcular las probabilidades solicitadas:

a) Probabilidad de que sea hombre, dado que tiene escolaridad de nivel secundaria.

$$P(H|S) = \frac{P(H \cap S)}{P(S)} = \frac{28}{78} = 0.35$$

La probabilidad de que sea hombre, dado que tiene escolaridad secundaria es de 0.35.

b) Probabilidad de que sea mujer, dado que tiene escolaridad de nivel bachillerato.

$$P(M|B) = \frac{P(M \cap B)}{P(B)} = \frac{17}{39} = 0.43$$

La probabilidad de que sea Mujer, dado que tiene escolaridad bachillerato es de 0.43.

6. La probabilidad de que un doctor diagnostique en forma correcta una enfermedad es del 70%. Cuando el doctor hace un diagnóstico incorrecto, la probabilidad de que un paciente presente una demanda es del 90%, ¿cuál es la probabilidad de que el doctor haga un diagnóstico incorrecto y el paciente presente una demanda?

Respuesta:

Como en todo ejercicio, debemos plantear las probabilidades para iniciar con la respuesta a este ejercicio.

$$P(C) = 0.70$$

$$P(I) = 1 - P(\mathrm{C}) = 0.30$$

$$P(D|I) = 0.90$$

¿Cuál es la probabilidad de que el doctor haga un diagnóstico incorrecto y el paciente presente una demanda?

$$P(\ I \cap D\ ) = P(D|I) \text{ x } P(I) = 0.90 \text{ x } 0.30 = 0.27$$

La probabilidad de que el doctor haga un diagnóstico incorrecto y el paciente presente una demanda es de 0.27

7. De la población estudiantil de cierta facultad, el 4% de los hombres y el 1% de las mujeres miden más de 180 cm. Además, el 60% de los estudiantes son mujeres. Ahora bien, si se selecciona al azar un estudiante y es de una estatura mayor de 180 cm, ¿cuál es la probabilidad de que se haya elegido a una mujer?

Respuesta:

Acomodamos nuestras probabilidades:

$$P(M) = 0.60$$
$$P(H) = 0.40$$
$$P(A|M) = 0.01$$
$$P(A|H) = 0.04$$

Para calcular dicha probabilidad usaremos el teorema de Bayes:

$$P(M \mid A) = \frac{P(M) \times P(A|M)}{P(M) \times P(A \mid M) + P(H) \times P(A \mid H)}$$

Por lo tanto:

$$P(M|A) = \frac{0.60 \times 0.01}{0.60 \times 0.01 + 0.40 \times 0.04}$$

La probabilidad de que se haya elegido a una mujer es de:

$$P(M|A) = 0.2727$$

8. Se dispone de dos métodos para transmitir un mensaje. El método A transmite correctamente el 70% de los mensajes y el método B el 90%. Un día se elige un método al azar y se transmiten ocho mensajes comprobándose posteriormente que los dos primeros se han transmitido de manera incorrecta. ¿Cuál es la probabilidad de que se haya utilizado el método A? ¿Cuál es la probabilidad de que se utilizara el método B?

Respuesta:

Como en todos los ejercicios, se debe indicar las probabilidades para iniciar con la respuesta a dicho ejercicio.

$$P(A) = 0.50$$
$$P(B) = 0.50$$
$$P(C|A) = 0.70$$
$$P(C|B) = 0.90$$
$$P(E|A) = 0.30$$
$$P(E|B) = 0.10$$

Al trasmitirse dos mensajes seguidos incorrectos, decimos que:

$$P(E2|A) = (0.30)2 = 0.09$$

$$P(E2|B) = (0.10)2 = 0.01$$

Aplicando el teorema de Bayes:

$$P(A|E^2) = \frac{P(E^2|A) \times P(A)}{P(E^2|A) \times P(A) + P(E^2|B) \times P(B)}$$

$$(B|E^2) = \frac{P(E^2|B) \times P(B)}{P(E^2|A) \times P(A) + P(E^2|B) \times P(B)}$$

Se sustituyen los valores:

$$(A|E^2) = \frac{0.09(0.50)}{0.09 \times 0.50 + 0.01 \times 0.50} = 0.90$$

$$(B|E^2) = \frac{0.01(0.5)}{0.09 \times 0.5 + 0.01 \times 0.5} = 0.1$$

De esta forma, la probabilidad de que se usara el método A es de 90% y la probabilidad de que se usara el método B es del 10%.

Agradecemos la participación en este proyecto al personal docente y estudiantes de la Facultad de Economía y de la Facultad de Psicología y al Instituto de Física.

Este material didáctico fue elaborado gracias al apoyo del Programa UNAM-DGAPA- PAPIME PE306125 y a las facilidades que el Programa Único de las Especializaciones en Economía de la División de Estudios de Posgrado de la Facultad de Economía, otorgó a los colaboradores del proyecto.

Actividad realizada con fines educativos.