

# La Investigación Cuantitativa en las Ciencias Sociales

2<sup>da</sup> Edición

**Autores:**

- Calizaya López José Manuel
- Alemán Vilca Yaneth
- Bellido Medina Rildo Santos
- Ceballos Bejarano Ferdinand Eddington



## Los Autores



**Calizaya López, José Manuel**, Doctor en Ciencias de la Educación, Magíster en Ciencias con Mención en Gestión Social - Desarrollo Sostenible. Licenciado en Trabajo Social. Docente Principal y Director del Departamento Académico de Trabajo Social, Docente investigador UNSA-INVESTIGA y miembro del Instituto de Investigaciones Sociales de la Universidad Nacional de San Agustín, Arequipa, Perú.



**Alemán Vilca, Yaneth**, Doctora en Administración (DBA), Maestra en Ciencias, con mención en Gerencia Social y de Recursos Humanos. Asesora de Proyectos de Investigación relacionados a la salud pública y mental. Docente de la Escuela Profesional de Trabajo Social de la Universidad Nacional de San Agustín



**Bellido Medina, Rildo Santos**, Doctor en Psicología y Magíster en Gerencia Estratégica de Recursos Humanos. Docente auxiliar en el programa de estudios de Relaciones Industriales, Docente investigador UNSA- INVESTIGA y miembro del Instituto de Investigaciones Sociales de la Universidad Nacional de San Agustín, Arequipa, Perú



**Ceballos Bejarano, Ferdinand Eddington**, Licenciado en Matemática, Magíster en Finanzas y Administración de Negocios, Doctor en Ciencias Empresariales. Docente auxiliar en el programa de estudios de Administración y Finanzas, y miembro del Instituto de Investigaciones Sociales de la Universidad Nacional de San Agustín, Arequipa, Perú



# La investigación cuantitativa en las ciencias sociales



La investigación cuantitativa en las ciencias sociales.

© Calizaya López José Manuel, Alemán Vilca Yaneth, Bellido Medina Rildo Santos, Ceballos Bejarano Ferdinand Eddington.

Diseño de portada: Manuelis Salazar

Diagramación: Manuelis Salazar.

Segunda edición Quito-Ecuador 2022 Reservados todos los derechos.

ISBN: 978-9942-42-027-5



No se permite la reproducción total o parcial de esta obra, ni su incorporación a un sistema informático, ni su transmisión en cualquier forma o por cualquier medio (electrónico, mecánico, fotocopia, grabación u otros) sin autorización previa y por escrito de los titulares del copyright. La infracción de dichos derechos puede constituir un delito contra la propiedad intelectual.

**Previa a su publicación, esta obra fue evaluada bajo la modalidad de revisión por pares anónimos**



# Contenido

Pag

CAPITULO 1.....	13
La investigación social cuantitativa.....	13
Introducción.....	13
Conociendo el tema de investigación.....	15
Descubriendo la problemática de la investigación.....	15
Elaborando un proyecto de investigación.....	16
Ideas y planteamiento de un problema de investigación.....	16
El problema de investigación como una o más preguntas.....	17
El problema como hipótesis a verificar o refutar.....	17
El problema como un sujeto de estudio.....	17
Como un propósito de acción.....	17
Revisión del soporte teórico, bases y fundamentos: Antecedentes del problema.....	18
Alcance del proyecto: Objetivos de la investigación.....	18
Diseño de la metodología.....	19
Diagrama de trabajo: recolección, procesamiento y análisis de datos.....	19
Elaboración del reporte final.....	20
Haciendo los antecedentes.....	21
Ejecución del proyecto de investigación.....	22
CAPITULO 2.....	25
Objetivos de investigación.....	25
Formulación de los objetivos de Investigación.....	26
Los objetivos se organizan en general y específicos.....	26
Niveles de complejidad de los objetivos de investigación.....	27
Redacción de los objetivos.....	27
Elementos constitutivos de la metodología de la investigación actual.....	30
Variables.....	30
Clasificación de variables.....	30
Variables cualitativas y cuantitativas.....	31



VARIABLES INDIVIDUALES Y COLECTIVAS.....	31
VARIABLES NOMINALES, ORDINALES Y PROPORCIONALES.....	31
Elaboración de constructos.....	32
Indicadores.....	32
Operacionalización de los indicadores.....	33
Índices.....	34
Hipótesis.....	34
Clasificación de las hipótesis.....	35
Hipótesis descriptivas.....	35
Hipótesis generales o teóricas.....	36
Hipótesis relacionales.....	36
Hipótesis direccionales y no direccionales.....	36
Hipótesis explicativas o causales.....	37
Hipótesis estadísticas.....	37
Características que debe tener una hipótesis.....	37
CAPITULO 3.....	39
La investigación cuantitativa y sus tipos.....	39
Etapas de una investigación cuantitativa.....	41
Diseño experimental en las investigaciones cuantitativas.....	42
Investigaciones cuasiexperimentales.....	47
Diseños no experimentales.....	50
Tipos de experimentos de diseños no experimentales.....	50
Encuestas según la temporalidad: seccionales y longitudinales.....	51
Encuestas seccionales.....	51
Encuestas longitudinales.....	51
Desarrollo del proceso al momento de aplicar una encuesta social.....	51
CAPÍTULO 4.....	54
Recolección de datos sociales.....	54
La encuesta Social.....	54
La encuesta descriptiva.....	55



Proceso de análisis de los datos de las encuestas descriptivas.....	57
La encuesta explicativa.....	58
La encuesta social.....	59
Determinación de la muestra.....	60
Población o universo.....	60
Marco de muestreo.....	60
Muestra.....	60
Unidad de muestreo.....	97
Fracción de muestreo.....	61
Representatividad de la muestra.....	61
Tipos de muestras.....	61
Tipos de muestras probabilísticas.....	61
Muestra aleatoria simple.....	61
Muestra estratificada proporcional.....	62
Muestra sistemática.....	62
Muestra de conglomerados.....	62
Tamaño de la muestra.....	63
Muestras no probabilísticas.....	64
El cuestionario.....	65
La escala.....	69
Tipos de escala.....	69
Escala por intervalos.....	69
Escala de razón.....	71
El inventario.....	71
Pruebas proyectivas.....	72
CAPÍTULO 5.....	73
Tratamiento de variables.....	73
(Análisis Descriptivo).....	73
Distribución de frecuencias absolutas.....	74
Distribución de frecuencias relativas.....	75
Distribución de frecuencias acumuladas.....	76



Presentación de resultados del tratamiento de variables.....	77
Medidas de tendencia central.....	81
Moda.....	81
Mediana.....	83
Media aritmética.....	85
Medidas de dispersión.....	88
La Varianza.....	88
Desviación estándar.....	89
Coefficiente de variación.....	90
Medidas de concentración.....	90
El Índice y Coeficiente de Gini.....	90
Quintiles.....	93
CAPÍTULO 6.....	94
Asociaciones y correlaciones Análisis inferencial.....	94
Comparación de muestras.....	95
Comparación entre dos proporciones.....	96
Comparación entre porcentajes.....	98
Comparaciones Múltiples.....	99
Análisis de varianza.....	102
Análisis de varianza simple.....	102
Análisis múltiple de varianza.....	106
Coefficientes de correlación para variable ordinales.....	107
Método de pares de valores.....	107
Coefficiente tau-a de Kendall.....	108
Coefficiente Q (de Yule).....	108
Coefficiente gamma de Goodman y Kruskal.....	109
Coefficiente rho de Spearman.....	110
Análisis de regresión.....	111
La Regresión Lineal Simple.....	111
Modelo de Regresión Lineal Múltiple o Polinómico.....	114
Modelos de Regresión Polinómicos (RP).....	115



Análisis multivariado.....	115
Análisis factorial.....	116
Análisis de clúster.....	116
Análisis de regresión.....	116
Análisis de desviaciones.....	116
Análisis discriminante.....	116
CAPÍTULO 7.....	117
Comunicación de la investigación cuantitativa.....	117
Artículos Científicos.....	118
Conferencias en congresos.....	119
Libros profesionales.....	120
Póster científico.....	121
CAPÍTULO 8.....	125
Validez y Confiabilidad de Instrumentos de Investigación.....	126
Validez.....	126
Validez de Constructo.....	128
Validez Predictiva o de Criterio Externo o Empírica.....	129
Cualidades deseables en una medida criterio.....	129
Confiabilidad.....	130
Coefficiente Alfa de Cronbach.....	131



# Prólogo

A lo largo de estos últimos cinco años los autores del libro *La investigación cuantitativa en las ciencias sociales*, han desarrollado competencias en la investigación asumiendo un rol académico fundamental en la formación de nuevos investigadores que contribuyan en el desarrollo de las sociedades.

La investigación cuantitativa es un método estructurado y sistematizado de recopilación y análisis de la información a través de diferentes fuentes, técnicas y herramientas con el propósito de cuantificar el problema de investigación, así mismo, se busca medir un fenómeno, ponderarlo y expresarlo en cifras, es así, como este tipo de investigación ayuda a identificar tendencias y niveles, realizar predicciones, comprobar hipótesis y obtener resultados generales en poblaciones grandes y diversificadas. Siendo un reto para las ciencias sociales encontrar la magnitud del problema y establecer su medición.

Por ello, el libro para esta segunda edición utiliza los conceptos claros y una forma didáctica para comprender la investigación social cuantitativa, desde la selección del tema, la identificación de la problemática social, los objetivos, las hipótesis, los tipos y diseños, los instrumentos, hasta el tratamiento de las variables integrando aspectos de la estadística que son relevantes para el análisis correspondiente. Además, de la importancia de comunicar la investigación a través de artículos científicos, conferencias en congresos, libros profesionales y póster científico.

Un aspecto central del libro es la utilidad del análisis estadístico tanto descriptivo como inferencial, resaltando la experiencia de los autores que de manera práctica te involucran en este proceso y motivan a profundizar en el mismo, sabiendo que es uno de los componentes más temidos entre los estudiantes e investigadores.

Así también, se evidencia la dedicación puesta en la redacción del manuscrito que invita a leerlo y a poner en práctica los conocimientos para la formulación del proyecto de investigación, siendo de mucha utilidad sobre todo para los estudiantes de pregrado y posgrado interesados en desarrollar sus competencias y actitudes hacia la investigación.

Para concluir felicito a las personas e instituciones que pudieron contribuir para hacer realidad esta obra, a la vez, agradezco de manera personal a los autores por invitarme a escribir algunas líneas que solo buscan valorar y reconocer el esfuerzo que la escritura de una obra científica tiene.

**Marcio Soto Añari, PhD**

Este libro es el resultado de varias investigaciones científicas de sus autores, en el área de las ciencias sociales, que aportan de manera significativa a la academia, la enseñanza y la educación en general.



# CAPÍTULO 1

## La investigación social cuantitativa

### Introducción

La investigación científica parte de un precepto fundamental que es la existencia de una estructura lógica, y por ende no contradictoria, que articula de manera fluida a todo el conocimiento humano que ha sido verificado, racionalizado y conceptualizado en un área determinada; a esto se le denomina ciencia. Cuando se extiende el concepto de ciencia a todos los campos de interés científico, entonces se está en presencia de las ciencias. Estas se definen según el campo de estudio que abordan; por ejemplo, las ciencias naturales o básicas como la Biología, la Química o la Física, se estudia a la naturaleza en general desde sus fundamentos elementales. Aquellas ciencias como la Astronomía, que hacen uso de diversas herramientas de otras ciencias en sus investigaciones, se les denomina ciencias aplicadas. Las ciencias que estudian al ser humano y su comportamiento se conocen como ciencias sociales y en este libro se abordará una metodología para desarrollar investigaciones científicas sobre estas últimas.

Ninguna investigación científica puede hacerse sin las personas que dedican su tiempo y esfuerzo a la confirmación o refutación de teorías científicas, a la búsqueda de nuevos conocimientos o a la ampliación de los ya existentes; a estas personas se les suele conocer como investigadores o científicos. Los investigadores pueden apoyarse, según sus propósitos en alguno de los dos, o en ambos (enfoque mixto) enfoques clásicos a conocer: el enfoque cualitativo y el enfoque cuantitativo.

En el enfoque cualitativo, conocido también enfoque fenomenológico trata de estudiar un problema sin extraerlo de su entorno, analizando sus interacciones para buscar patrones que modelen su comportamiento. En las ciencias sociales la metodología cualitativa toma como modelos la teoría crítica que estudia las estructuras sociales imbricadas como la política, los grupos étnicos y la economía entre otros, y el constructivismo que busca reconstruir el camino que modeló al problema estudiado.

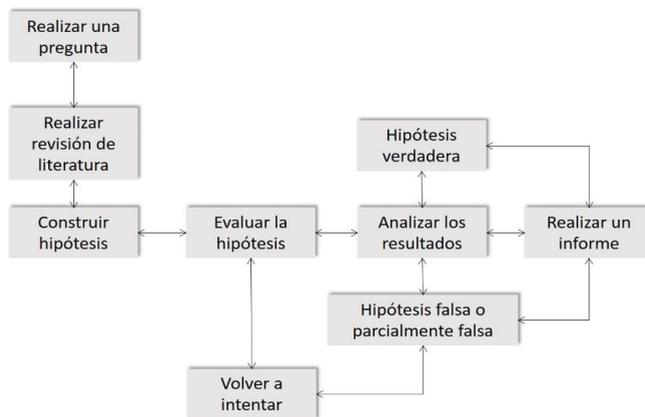
En el enfoque cualitativo el investigador se vuelve parte del problema y en muchas ocasiones pierde la objetividad.

El enfoque cuantitativo por otro lado parte del paradigma explicativo del positivismo, donde el acercamiento a la investigación de un problema se realiza desde un punto de vista racional, objetivo, que permite la observación, manipulación y verificación de predicciones acerca del mismo. La verificación de postulados e hipótesis muestran la validez o falsedad de los mismos, por lo tanto, cualquier problema planteado debe ser susceptible a comprobaciones experimentales que prueben su validez y de esta manera pueda ser aceptado como parte constituyente de una teoría formal

En el enfoque cuantitativo prevalece la subjetividad, el investigador no es parte del problema y las perspectivas se hacen en función de los datos observados.

El enfoque cualitativo tiene su origen de las investigaciones científicas de las ciencias naturales, que luego fueron adoptadas por las ciencias sociales con bastante éxito y han demostrado además ser de mucha utilidad en esta área. Esta es la metodología que se usará en este trabajo y a partir de ella se desarrollarán las directrices que son necesarias para la ejecución de cualquier proyecto de investigación científica en ciencias sociales.

Los investigadores que utilizan la metodología de investigación cuantitativa hacen uso de un modelo o estrategia general que les provee las herramientas adecuadas para ello, a este modelo se le conoce como el método científico, el cual consiste en una serie de etapas que permiten elaborar conclusiones sobre aquello investigado. Por lo general el método científico consta de las siguientes etapas: 1) Formulación de preguntas que definan un problema 2) Investigar a profundidad el problema 3) formulación de una hipótesis 4) verificación o refutación de la hipótesis 5) Obtención de resultados 6) Generación de conclusiones 7) divulgación o socialización del trabajo realizado. En la Fig. 1. se puede observar de manera resumida el proceso que involucra al método científico.



**Figura 1. Esquema resumido del método científico**

El método científico es en sí una metodología, una forma estandarizada que sigue el camino del rigor científico y por ello requiere que los pasos antes descritos se realicen con especial dedicación y cuidado. Sin embargo, las etapas antes descritas no muestran cómo los investigadores deben o pueden abordar un problema a estudiar, no les ayuda en la formulación de preguntas, de construcción de experiencias que permitan la validación de propuestas. Por su carácter general ha sido necesario el desarrollo de mecanismos que esclarezcan la manera en la que los investigadores pueden acercarse a un problema y abordarlo científicamente y es este el propósito de esta sección.

Teniendo en cuenta la información hasta ahora expuesta se puede concluir que toda investigación científica posee el siguiente desarrollo general.

- a) Un primer acercamiento al problema que se quiere abordar el cual parte de una idea de investigación.
- b) La preparación de un plan o proyecto de investigación donde se indica el proceso a seguir para abordar el problema.

- c) Se desarrolla el proceso de ejecución del plan. Esto es la instrumentalización del plan o proyecto.
- d) La socialización de los resultados obtenidos.

### **Conociendo el tema de investigación**

Como se ha mencionado en un párrafo anterior, en una investigación científica, los investigadores buscan verificar, rebatir, ampliar o proponer nuevos aportes experimentales o teóricos al conjunto de la ciencia que están abordando. Cuando se inicia una investigación científica se hace desde un punto de vista original del investigador sobre un problema, o a partir de los aportes que otros investigadores han hecho al tema y donde esos aportes presentan un mayor o menor grado de profundidad y validación.

Esta premisa implica al investigador estar inmerso en un proceso continuo de actualización en lo que respecta a los aportes que se están haciendo a esa línea de conocimiento particular. El investigador a menudo desarrolla su trabajo partiendo de su formación teórica y metodológica, por lo que sus inquietudes investigativas lo direccionan hacia la ampliación del conocimiento partiendo de una línea de investigación, la cual también le une a una comunidad de investigadores con inclinaciones y temas investigativos similares.

La investigación científica no se establece solo entre investigadores experimentados, sino que da cabida a nuevos entusiastas de la ciencia y el conocimiento. Para aquellos investigadores que se inician en el proceso de creación o ampliación de conocimiento científico se recomienda un grado de inmersión bibliográfica más profunda como también la guía y apoyo de investigadores que, con una formación consolidada pueden servir de mentores y apoyar su proceso de crecimiento en el área. En última instancia se recomienda una aproximación gradual a los temas de investigación que deseen desarrollar a fin de que las competencias requeridas para realizar trabajos de calidad científica estén acorde al nivel del conocimiento que domina el nuevo investigador.

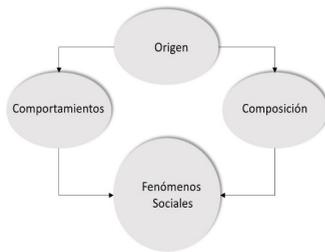
### **Descubriendo la problemática de la investigación**

El acercamiento a un problema de investigación requiere en un principio de una evaluación metodológica, de pensar cómo abordar el problema en sí para definir qué estrategias y recursos deben ser utilizados a fin de poder estudiarlo de manera satisfactoria. En este libro se presentará cómo aproximarse a un problema de interés social utilizando la metodología de investigación cuantitativa. El desarrollo de cualquier problemática en este campo se realizará mediante esta metodología.

La investigación social puede desarrollarse considerando el método científico.

La investigación social cuantitativa nace de las investigaciones y aportes teóricos de Auguste Comte y Emile Durkheim, los padres del positivismo, y quienes concluyeron que los fenómenos sociales pueden ser estudiados utilizando el mismo método que utilizan las ciencias naturales, que es el ya mencionado y descrito método científico.

Sin embargo, en la actualidad el enfoque cuantitativo para las ciencias sociales está más adscrito a las corrientes pospositivistas que permiten una mayor flexibilidad al acercarse a un problema concreto. En este último enfoque se pasa del conocimiento como copia de la realidad objetiva, a la del conocimiento como producto de la interacción que existe entre el conocedor y el sujeto o problema conocido.



### **Elaborando un proyecto de investigación**

Un proyecto de investigación es un trabajo documentado donde se plasma el proceso que inicia con el acercamiento al problema que se desea investigar hasta los resultados obtenidos de las experiencias elaboradas para tal fin. El proceso comprende una serie de etapas que registran de qué manera se trató el problema y cuáles fueron sus resultados y consecuencias.

Los componentes o etapas de un proyecto de investigación social cuantitativo se pueden mostrar de la siguiente manera:

- a) Ideas y planteamiento de un problema de investigación
- b) Revisión del soporte teórico, bases y fundamentos: Antecedentes del problema.
- c) Alcance del proyecto: Objetivos de la investigación
- d) Diseño de la metodología
- e) Diagrama de trabajo: recolección, procesamiento y análisis de datos.
- f) Elaboración del reporte final.

### **Ideas y planteamiento de un problema de investigación**

Todo proyecto de investigación social inicia con una idea de investigación, que tiene como propósito

el encontrar una respuesta o solución a un problema que permita el avance, la ampliación, la refutación o verificación de un conocimiento en las ciencias sociales. La idea de investigación indica qué es aquello que se busca investigar y frecuentemente engloba un conjunto de temas que luego van siendo acotados hasta llegar a la identificación del problema a ser abordado para la búsqueda de su solución.

Los problemas que abordan la investigación en ciencias sociales pueden ir desde una situación social determinada como las ventajas y desventajas de la inmigración venezolana en Colombia, la optimización del aprendizaje bajo condiciones de pandemia en Perú, o pueden ser acercamientos teóricos como, por ejemplo, cuáles factores eternizan el populismo en América Latina, la precariedad de los sindicatos, la debilidad de las democracias modernas, la brecha digital en América, entre otros.

La fuente de ideas para un planteamiento investigativo pueden partir desde un acercamiento tanto práctico como teórico, sin embargo, con independencia de su origen, el problema social a investigar puede plantearse como: una o varias preguntas, una hipótesis a verificar o refutar, como un sujeto de investigación o como un propósito de acción.

### **El problema de investigación como una o más preguntas**

Supongamos que la investigación se centra en nivel de destreza en matemáticas de alumnos de primer nivel en una región. Se puede plantear un problema de investigación diseñando las siguientes preguntas:

¿Incidirá el nivel educativo del padre o la madre en las destrezas del alumno?

¿Dónde existe un mayor nivel de destrezas en matemáticas de los alumnos, en las escuelas que utilizan el método de enseñanza tradicional o las que se apoyan en nuevas metodologías?

### **El problema como una hipótesis a verificar o refutar**

Aquí se plantea como la verificación o refutación de una sentencia. Por ejemplo, Comprobar si los alumnos cuyos padres poseen un nivel educativo superior tienen mayores destrezas matemáticas. Comprobar si los alumnos que aprenden matemáticas mediante metodologías activas de aprendizaje poseen más destrezas que aquellos alumnos que aprenden matemáticas mediante el método tradicional.

### **El problema como un sujeto de estudio**

Se ejecuta a manera de proposiciones que pueden o no estar alineadas con preguntas iniciales relacionadas a un problema. Las proposiciones pueden tener la siguiente enunciación.

“El problema a abordar en la investigación se divide en: 1) análisis de las destrezas en matemáticas de los alumnos de padres con educación superior en comparación de aquellos alumnos que no cuentan con padres con esa formación. 2) Estudio de las destrezas en matemáticas en los alumnos que siguen el modelo tradicional de aprendizaje en comparación con los que siguen nuevas metodologías.

### **Como un propósito de acción**

Se busca que el problema o los problemas en los que se fundamenta el proyecto de investigación se



redacte como una propuesta de mejora. Podría redactarse de la siguiente manera:

“Cómo mejorar el nivel de destrezas en matemáticas de los alumnos de primer nivel de la región X: propuesta metodológica”

### **Revisión del soporte teórico, bases y fundamentos: Antecedentes del problema.**

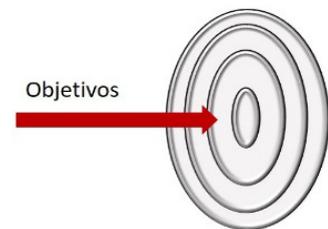
Cuando ya se ha planteado un problema que será el eje del proyecto de investigación, la etapa siguiente es la sustentación teórica del mismo. Para ello se inicia una revisión metódica de la literatura científica que se ha desarrollado entorno al problema a estudiar, esta revisión incluye las teorías, las experiencias desarrolladas y relacionadas al tema y las investigaciones en general que, una vez evaluados y procesados, definirían los antecedentes válidos del proyecto. Esto le brinda al mismo una sustentación conceptual y metodológica que ayudaría al investigador a orientar y darle contexto a su investigación. Entonces, de lo antes descrito podemos deducir que, la revisión de la literatura nos brinda el marco teórico dentro del cual se encuentra el proyecto. La labor de síntesis del marco teórico nos brinda los antecedentes del proyecto.

Los antecedentes tienen la importancia de ubicar la investigación dentro del campo al cual pertenece, destacando los aportes que otros investigadores han alcanzado sobre el tema que se busca desarrollar. Los antecedentes además conforman las referencias más directas a los resultados sobre la misma línea investigativa a la que pertenece el proyecto de investigación.

### **Alcance del proyecto: Objetivos de la investigación**

Alcance del proyecto: Objetivos de la investigación  
Una vez identificado el problema que se va a abordar en el proyecto de investigación, se deben definir los objetivos del mismo, que pueden clasificarse en generales o específicos, y representan a aquello que se pretende resolver, la manera de lograrlo y los resultados que se aspiran alcanzar. La identificación del o los objetivos del proyecto ayuda a orientar y definir las etapas que requiere el proyecto, los propósitos de cada una de ellas y además a conocer el alcance, la amplitud y el rango que tendrán. Los objetivos deben poseer claridad y precisión, por lo tanto, deben responder a los elementos constitutivos básicos del problema tratado; en la metodología cuantitativa, los objetivos deben tener la capacidad de ser medidos, observados y verificados, deben al-

Los objetivos deben ser claros y concisos, son el eje de la investigación y deben conducirla a los resultados.



**Figura 3. Los objetivos deben ser concisos.**

canzarse mediante un proceso secuencial lógico en cada una de las etapas diseñadas para tal fin. Entonces los objetivos proveen una guía definida para la recolección e interpretación de las muestras o datos que son necesarios para alcanzar las metas propuestas. Una vez comprobados los resultados, estos serán añadidos al corpus teórico del que forma parte como un nuevo aporte original, ampliando el conocimiento en esa rama de la ciencia. Los objetivos deben poseer claridad y precisión, por lo tanto, deben responder a los elementos.

### **Diseño de la metodología**

Con los objetivos definidos, el investigador debe elaborar la metodología necesaria para resolver el problema o los problemas planteados. Dentro de la metodología destacan las estrategias y las técnicas que le permitirán alcanzar dichos objetivos, comprobando y verificando si coinciden con los supuestos teóricos esperados.

El diseño metodológico contempla la manera adecuada de recolectar las muestras o datos necesarios para llevar a cabo la investigación, especificando los procedimientos con detalle a fin de no introducir anomalías en los mismos.

### **Diagrama de trabajo: recolección, procesamiento y análisis de datos**

Una vez concebida la aproximación metodológica que dará respuestas al problema planteado según los objetivos definidos, toca el momento de desarrollar las etapas de recolección, procesamiento y análisis de la información, la misma que estará constituida por las muestras o datos que se utilizaran para tal fin.

a. **Recolección de datos:** para recabar las muestras o datos necesarios para el desarrollo de la investigación se requiere del uso de las estrategias concebidas en el diseño metodológico, que pueden incluir entrevistas, cuestionarios, encuestas, material documental, estadísticas entre otros. Los datos se recolectan según mecanismos establecidos donde se detallan los procedimientos utilizados, además de información como: puntos de recolección, grupos etarios, estratos socioeconómicos, fuentes, permisos, consentimientos, entre otros. Es necesario resaltar que se debe realizar un trabajo en dos direcciones para que la recolección de información sea exitosa. Por una parte, el personal que se destinará en la recolección de los datos debe estar entrenada para ello, deben contar con el material suficiente para realizar su labor sin contratiempos y tener claras las estrategias que han de utilizar. Por otra parte, se debe tener clara la población foco del trabajo y preparar el acercamiento de tal manera que la dinámica necesaria para la investigación no se resienta por algún contratiempo, evitando dificultades en la recolección de muestras. Cuando los datos tengan carácter documental es necesario la elaboración de hojas de ruta donde se especifiquen los documentos que deben ser consultados y la información que debe ser registrada.

La recolección de los datos se lleva a cabo siguiendo un plan preestablecido donde se especifican los procedimientos para la recolección, incluyendo la ubicación de las fuentes de información o los sujetos, el lugar de aplicación, el consentimiento informado y la manera de abordarlos.

b. **Procesamiento de datos:** Al ser recolectados los datos que serán necesarios para la instrumenta-

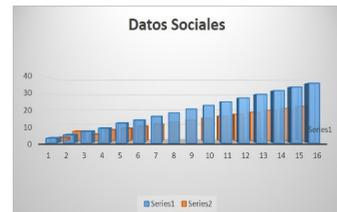


ción del proyecto se procede a su normalización que consiste en aplicarles un conjunto de procesos de depuración que permitan su posterior uso. Es lo que se conoce como darles formato a los datos. Lo primero que se debe hacer es examinar cada conjunto de datos según su naturaleza, observando si existen incongruencias, inconsistencias, errores u omisiones que deban corregirse si es posible, de no serlo se procederá a depurar la data, eliminando aquellos elementos que no cumplan los estándares buscados.

El proceso de formateado de los datos consiste en codificar los datos según un patrón de referencia establecido que permita una manipulación eficaz de los mismos; es similar a, por ejemplo, si se requiere transportar calculadoras de un punto a otro, la mejor manera de hacerlo es colocando una cantidad específica de calculadoras de un punto a otro, la mejor manera de hacerlo es colocando una cantidad específica de calculadoras dentro de una caja diseñada para alojar un número exacto de ellas, de esta manera es más sencillo contabilizarlas y transportarlas. De la misma manera los datos se organizan dentro de un sistema para una eficaz manipulación de los mismos. En las investigaciones cuantitativas es necesario que los datos puedan ser mensurables y proclives a ser analizados de manera matemática, por tanto, uno de los primeros pasos es vaciar los datos en formatos digitales como hojas de cálculo y bases de datos, y asignarles valores numéricos según tablas convertibles y la creación de categorías según los tipos de datos.

El análisis de datos debe comprender una visión matemática de la información y debe focalizarse en los objetivos del planteamiento del problema.

c. Análisis de datos: Con la data depurada, organizada y dentro de un formato que permita su comprensión y manipulación, se entra en la fase de análisis de la información resultante. Este análisis llevara un orden lógico que permita encontrar tendencias, patrones y/o relaciones significativas en la información. El análisis busca identificar las unidades básicas irreductibles de las que se compone la información para conseguir en ellas las razones del por qué se relacionan de esa manera y no de otra forma. Esto permite hacer deducciones y conclusiones que sintetizen el resultado obtenido.



Las investigaciones cuantitativas utilizan como técnica principal el análisis estadístico para el procesamiento de información. El análisis estadístico permite la representación de la información mediante porcentajes, también permite identificar la desviación media, identificar correlaciones, identificar variables a fin de realizar comparaciones entre los resultados obtenidos y los esperados teóricamente. La contrastación entre los resultados esperados y obtenidos son quienes muestran el grado de éxito del proyecto de investigación. Por lo general los análisis estadísticos se realizan a través de programas que facilitan la ejecución de complejas operaciones matemáticas en tiempos mucho menores que los que le tomaría a un investigador realizarlos manualmente.

## Elaboración del reporte final

Ya en esta etapa es cuando se procede a la divulgación de los resultados obtenidos en la ejecución del proyecto de investigación. En este punto ya se ha procedido al análisis e interpretación de los resultados que arrojaron la investigación. Ya se cotejaron los resultados experimentales con los objetivos buscados, verificándose si los postulados o las hipótesis se cumplieron o no y qué consecuencias generan dichos resultados, cómo se relacionan con el trabajo de otros investigadores y que aportan al acervo científico del área investigada.

Es importante que los proyectos científicos desarrollados en cualquier área sean divulgados pues es la manera más importante de ayudar a extender los conocimientos que se tienen sobre un área del conocimiento específica. La publicación de las investigaciones científicas ayuda a otros investigadores a formular nuevas preguntas, a encontrar posibles soluciones a problemas no estudiados apoyándose en los resultados de los trabajos ya publicados y que les sirven como referencia para sus investigaciones.

Las investigaciones científicas pueden publicarse de diferentes maneras: mediante artículos para ser publicados en revistas indizadas, mediante disertaciones en conferencias, mediante exposiciones en congresos o ponencias.

**Haciendo los antecedentes** :Los antecedentes son aquellos trabajos anteriores (que pueden ser artículos científicos avalados por revistas indizadas, tesis de grado o posgrado, resultados de investigaciones institucionales de carácter público, conferencias, congresos entre otros.) que se han ejecutado en la misma área investigativa y que poseen vínculos temáticos directos con el proyecto que se está desarrollando, por tanto, esas referencias representan las bases procedimentales y teóricas en las cuales este se apoya. Los antecedentes brindan un marco de referencial al lector para comprender qué se esta abordando en el proyecto, delimitar la importancia de la investigación, cuáles problemas aborda que no se hayan investigado o que se han investigado pero que necesitan ser corregidos, ampliados o refutados. Con ello se busca el fortalecimiento del conocimiento en el área de interés. Un ejemplo hipotético de antecedentes para un proyecto de investigación se muestra a continuación.

Tema: Formación docente en el manejo de TIC y su importancia en el proceso de enseñanza aprendizaje en la educación superior en Latinoamérica.

En el siglo XXI, la masificación de las tecnologías de la comunicación y la información (TIC) han cambiado radicalmente la manera en la que consumimos y procesamos la información. La globalización de internet, la aparición de las redes sociales y la cultura digital ha impactado directamente sobre el comportamiento social a todo nivel e influye en todos los ámbitos en los cuales se desarrolla la actividad humana. La educación no iba a ser una excepción de esta realidad [1].

La socióloga Lucy Morningstar, en su tesis De TIC ni me acuerdo: una propuesta metodológica para eliminar el bajo nivel de competencias digitales del profesor universitario, realiza un análisis detallado del por qué los profesores universitarios en un muy importante porcentaje se resisten a la adopción de las TIC en sus sistemas de enseñanza aprendizaje y cómo esto les resta com-

petitividad a las universidades y cómo además impacta en la diversificación de la economía. El estudio contempla un acercamiento cuantitativo donde mediante la aplicación de encuestas, recolección de mallas curriculares, verificación de acondicionamiento de aulas, comprobaciones directas de competencias tecnológicas en docentes a través de ensayos prácticos se creará una métrica que, mediante las ecuaciones que modelen la tendencias en las diferentes variables identificadas se pueda aplicar acercamientos estadísticos-conductuales que permita definir una estrategia de menor resistencia a la adopción de nuevas tecnologías en los procesos de enseñanza aprendizaje de los profesores.

Trabajos como el de La Hoz, muestran que existe una tendencia a resistir el uso de nuevas tecnologías en aquellos profesores que se inclinan por enfoques conductistas centrados en el profesor, pero no muestra de qué manera esta tendencia pueda revertirse. Trabajos como los de Arena se centran en las estadísticas de uso de las nuevas tecnologías, pero no muestran como impactan el uso de la tecnología en la competitividad de las universidades ni en la capacidad potencial de ayudar a los alumnos a alcanzar competencias y habilidades que le permitan manejar perspectivas económicas mejores.

En Latinoamérica un importante número de docentes se resisten al uso de nuevas tecnologías y utilizar las computadoras como herramientas complementarias de la enseñanza.

El uso de las TIC en la educación puede representar una herramienta de gran relevancia para contribuir en el proceso de enseñanza de forma significativa.

El trabajo de Morningstar muestra que existen mecanismos que pueden incentivar a los profesores a adoptar metodologías que promuevan las TIC no como mera herramienta complementaria, sino como soporte del proceso de enseñanza aprendizaje.

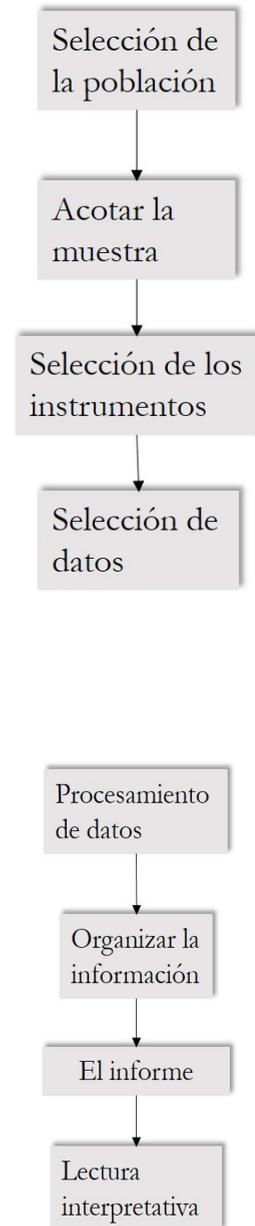
Mornigstar concluyó que, estructurando un perfil del docente que incluye, edad, género, formación docente, información sobre las TIC, salario, entre otras variables los mismos utilizaran con mayor o menor grado las TIC en los procesos de enseñanza aprendizaje que diseña para sus alumnos. También se estudiaron los factores externos al aula como las políticas institucionales, posiciones del plantel administrativo hacia las TIC, políticas nacionales y locales para conocer que tanto a favor o en contra, o incluso, si existe alguna realmente incidencia positiva o negativa en los casos concretos estudiados. Depurando todos estos datos se logró un enfoque matemático-procedimental que luego de varias pruebas arrojaron un avance positivo en la mejora tanto de la capacidad docente en el manejo de las TIC y su impacto en las habilidades asimiladas por los alumnos. Según lo cotejado un 53, 2% del alumnado se sintió más capacitados en las áreas de matemáticas e informática. Se estima que aún deben realizarse estudios complementarios para afinar el trabajo realizado por Morningstar.

### **Ejecución del proyecto de investigación**

Cuando ya está definido el proyecto de investigación en todas sus partes constituyentes se procede a su ejecución. En esta instancia se conoce cuál será el grupo de colaboradores que, junto con el investigador desarrollaran el proyecto, se establecen los roles y las actividades específicas según el diseño preestablecido de tal manera que las etapas del proyecto se cumplan de manera secuencial y oportuna. En este momento

se irán nombrado para en capítulos posteriores ampliarlas con más detalles. Las actividades y etapas pueden implementarse de la siguiente manera:

- Selección de la población objetivo, y si es pertinente se toma una muestra de la misma.
- Se acota el tamaño de la muestra y las características de la misma.
- Selección de los instrumentos de recolección de datos (encuesta, formularios, cuestionarios, etc.).
- Ensayos de efectividad de los instrumentos; en caso de introducir modificaciones se realizará el rediseño de los instrumentos a utilizar.
- Entrenamiento de los equipos de apoyo y de recolección de datos mediante pruebas de campo.
- Selección y depuración de datos (descartando datos incompletos, ambiguos, falsos, etc.).
- Estructuración de los datos depurados en un formato ya prediseñado para su manipulación.
- Acondicionamiento del modelo de análisis a utilizar según lo definido en los objetivos del proyecto (tablas estadísticas, histogramas, cálculos estadístico-descriptivo, regresiones, correlaciones, etc.).
- Elección de paquetes de procesamiento estadístico (Excel, SPSS, JASP, Jamovi, etc.). Ponderar el uso de procesos estadísticos de manera manual si la población y los datos no son muy grandes.
- Organización de la información procesada para la elaboración del informe.
- Elaboración de las partes del informe: introducción, antecedentes, capítulos, anexos, soportes.
- Preparación de la bibliografía pertinente al proyecto según las normas de citación establecidas previamente (APA, IEEE, Vancouver, etc.).



- Lectura interpretativa de los análisis hechos a la información obtenida.
- Primer borrador del informe, este debe contener:
  - Síntesis abreviada de la investigación, presentado los hallazgos o resultados más importantes del mismo.
  - Introducción al proyecto.
  - Marco teórico, conceptual o antecedentes según la naturaleza del proyecto.
  - Problema de investigación abordado y su envergadura (generales, específicos).
  - Modelo metodológico empleado.
  - Presentación de los resultados obtenidos y verificados.
  - Interpretación de los resultados en contraste con los supuestos teóricos y experimentales de referencia.
  - Limitaciones del proyecto
  - Presentación de conclusiones
  - Bibliografía de referencia general y específica relacionada directamente con el proyecto.

# CAPÍTULO 2

## Objetivos, Hipótesis y Variables

### Objetivos de investigación

La formulación de objetivos es una de las consecuencias más importantes que ha producido la incorporación de modalidades de investigación diferentes al enfoque positivista donde el papel confirmatorio o denegatorio de la hipótesis era central y alrededor de la cual giraba las labores de la investigación. Ahora, y ante la evidencia de necesidades investigativas que no requieren la formulación de una hipótesis, que como se indicó, reúne a las investigaciones confirmativas o denegativas, se han desarrollado modalidades de investigación que requieren de la construcción de objetivos de investigación para poder responder a sus propósitos. Entre los modelos metodológicos de investigación que hacen uso de la construcción de objetivos están las metodologías descriptivas, comparativas, proyectivas entre otras.

Los objetivos para una investigación científica tienen como orientación principal la obtención de conocimiento. Una investigación científica busca extender el conocimiento sobre un área dentro de la ciencia; sin embargo, es necesario diferenciar cuales son objetivos de investigación y cuáles no. Por ejemplo, una revisión de material bibliográfico sobre un tema que tiene como ejecutivo la recopilación y presentación de material relacionado con un tema no es un objetivo de investigación válido, pues no añade nuevo conocimiento sobre él. El siguiente enunciado no es un objetivo válido de investigación:

La indagación teórica es principalmente bibliográfica y la presentación de síntesis organizadas sobre los conceptos desarrollados por los autores citados no genera nuevo conocimiento. No sería material para una investigación científica sino propia de un trabajo monográfico.

Un ejemplo de enunciado de objetivo de investigación correcto se puede apreciar a continuación:

En el ejemplo anterior el investigador requiere aplicar al menos un instrumento que le permita recopilar información y con ella crear una estructura organizada que le provea información que sea cuantificable y medible y de ello analizar los resultados que arrojan.

“Ejecutar una investigación teórica que revise los acercamientos de diversos autores sobre el concepto de sindemia y su importancia en el contexto de la pandemia del covid-19”.

“Analizar los lineamientos de las políticas pública para la educación superior en el Perú desde las perspectivas de su pertinencia y calidad”.



Las conclusiones producto de este proceso serán un conocimiento distinto y que se añade al cuerpo teórico del área que se ha trabajado.

Existen también aquellos proyectos de investigación que son denominados como “especiales” este tipo de trabajos son aquellos que conducen a la creación de algo, pero que no son el resultado de un proceso investigativo.

## **Formulación de los objetivos de Investigación**

**Los objetivos se organizan en general y específicos.**

### **Objetivo general**

El objetivo general de un proyecto de investigación científica busca identificar qué es aquello que se pretende alcanzar durante el desarrollo de la investigación, esto significa que debe ser consistente con el problema y ser expresado de manera clara e inequívoca, enfocada en las metas a alcanzar. Dichas metas deben además ser importantes, realistas, medibles, sin ambigüedades y alcanzables dentro del proyecto investigación, de esta manera se evitan confusiones o desvíos en el desarrollo del mismo. Lo recomendado por todas las metodologías de investigación es que la investigación gire en torno a un solo objetivo general.

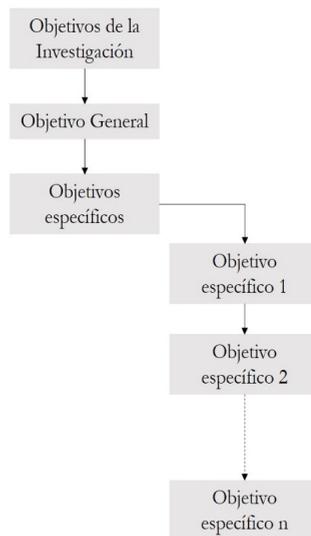
### **Elementos que deben constituir al objetivo general**

El objetivo general hace uso de verbos que indican el grado de complejidad que posee el problema a abordar. Ejemplo: demostrar, comparar, analizar, etc. Además, el objetivo general identifica quien ejecuta las etapas o subobjetivos que lo componen, declarando la secuencia en que dichas metas deben ser alcanzadas y el tiempo en el que deben alcanzarse.

El objetivo general es la guía de todo el proyecto, es necesario tenerlo siempre en cuenta a lo largo de las etapas que conforman al proyecto. Es probable que el objetivo general para poder ser abordado requiera ser subdividido en varios objetivos. A estos se les denomina objetivos específicos.

### **Objetivos específicos**

Los objetivos específicos se derivan del objetivo general y se formulan de tal manera que cumplan de manera diferenciada cada uno de los aspectos que constituyen el objetivo general; es decir, los objetivos específicos dan respuesta a las interrogantes que conforman al objetivo general, ellos responden a los qué, cuándo, cómo y dónde del problema a abordar. Para una mejor comprensión, si el objetivo general del proyecto de investigación es comparar los resultados con otros obtenidos en investigaciones anteriores, será necesario tener objetivos que describan las



acciones realizadas en una y otra investigación y otros objetivos que establezcan comparaciones entre ellas. Si el objetivo general es explicar un fenómeno o situación serán necesarios objetivos que describan, comparen, analicen y expliquen la situación o fenómeno. Puede suceder que el objetivo de investigación sea proponer un nuevo método para cierto proceso, se necesitaran objetivos específicos que describan, comparen, analicen, expliquen, predigan o proyecten ciertos resultados o procesos obtenidos de la investigación.

### **Niveles de complejidad de los objetivos de investigación**

Se entiende que el objetivo general representa el nivel más alto de complejidad en la elaboración del conocimiento. Entonces los objetivos específicos representan subniveles que deben cumplirse para lograr el objetivo general. Algunos autores han identificado hasta diez categorías en las cuales se pueden identificar los objetivos de investigación según su grado de complejidad. A su vez estos grados de complejidad pueden agruparse en cuatro niveles.

**Nivel perceptual:** donde el investigador busca acercarse a la problemática a estudiar destacando las características más evidentes o manifiestas del mismo, intentando apreciar el problema tal como se presenta, sin añadir puntos de vistas subjetivos. A este nivel pertenecen los objetivos de investigación tales como explorar y describir.

**Nivel aprehensivo:** en este nivel el investigador ofrece una reinterpretación de las observaciones que ha hecho a las características de una problemática a fin de brindar nuevos aportes sobre la comprensión de aspectos del problema que no fueron tan evidentes en un primer visionado. En este nivel se ubican los objetivos que buscan comparar y analizar.

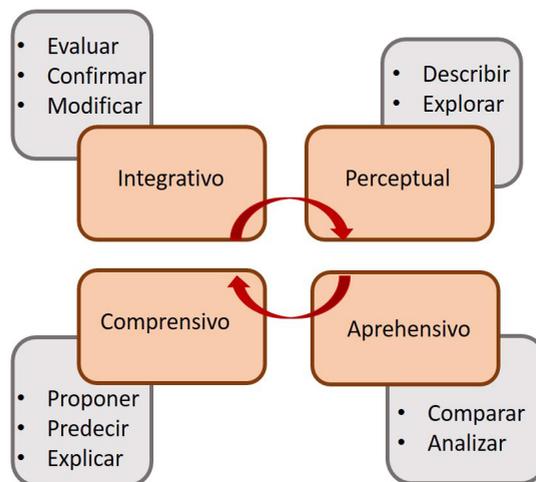
**Nivel comprensivo:** en este nivel el investigador busca explicar el por qué una problemática ocurre de cierta manera y como se relaciona con los diferentes aspectos que la componen. Las explicaciones permiten hacer predicciones sobre el problema y ofrecer en algunos casos información sobre cómo realizar manipulaciones, transformaciones y aplicaciones sobre lo investigado. En este nivel se ubican los objetivos que buscan explicar, predecir, proponer.

**Nivel Integrativo:** donde el investigador puede ir más allá del nivel comprensivo y proponer acciones que permitan obtener desde el problema estudiado resultados novedosos. En este nivel se ubican los objetivos que buscan modificar, confirmar y evaluar. Lo ante expresado se puede apreciar en la figura 3.

### **Redacción de los objetivos**

Los objetivos deben ser expresados de forma clara, ser concretos, observables, medibles, adecuados y realistas. Se ha adoptado por norma general que los objetivos se enuncian mediante verbos en infinitivo que expresen una acción que a su vez permite alcanzar una meta. Cada objetivo debe estar enfocado en una única meta y por tanto no en los medios o recursos involucrados en la consecución de la misma.

Un objetivo es claro cuando al leerlo una persona cualquiera puede comprender sin lugar a dudas aquello que se pretende resolver o conocer en el proyecto de investigación. Además, la misma persona puede comprender de qué manera se buscará resolver o conocer el problema y bajo qué pautas o condiciones se pretende hacer esto



**Figura 3. Niveles de complejidad de los Objetivos de investigación.**

La redacción de los objetivos debe ser capaz de enunciar a qué nivel se pretende alcanzar la meta planteada. Cuáles serán las unidades o criterios de medida que darán soporte a los resultados a fin de evidenciar el grado de desempeño o ejecución obtenidos.

Los objetivos de investigación tienen una relación directa con el problema de investigación que se está abordando. Esta relación puede ejemplizarse según la Tabla 1 donde dependiendo de la pregunta de investigación se elaboran las preguntas pertinentes.

**Tabla 1. Relación entre las preguntas que dirigen la investigación y el objetivo general.**

Posibles preguntas	→	Objetivo General
¿Qué quiero saber? ¿Qué hay?	→	Explorar
¿Cómo es?, ¿Cuántos casos...? ¿Qué intensidad tiene?	→	Desarrollar
¿Cómo se puede interpretar? ¿Cuáles son los significados menos evidentes? ¿En qué medida se corresponde?	→	Analizar
¿Qué diferencias o semejanzas hay entre estos grupos en relación a este evento?	→	Comparar
¿Por qué ocurre?, ¿Cuáles son las causas?	→	Explicar
¿Cómo se presentará este fenómeno en un futuro?	→	Predecir
¿Cuáles serán las características de un diseño?	→	Proyectiva
¿Qué cambios se pueden producir durante?	→	Interactiva
¿Existe relación entre estos eventos?	→	Confirmatoria
¿Cuál es la efectividad? ¿Está alcanzando los objetivos?	→	Evaluativa

Como se puede apreciar en la relación antes descrita podemos observar que los verbos en infinitivo pueden ser del tipo: representar, determinar, demostrar, examinar, especificar, indicar, analizar, estimar, comparar, valorar y relacionar respecto de los conceptos o variables incluidas.

Ejemplo de objetivo general y específicos

Veamos un ejemplo de cómo se redactan los objetivos, tanto generales como específicos.

Ejemplo 1:

**Título de la investigación:**

Actitud de los directivos nacionales respecto de la innovación y la creatividad en sus empresas.

**Objetivos de la investigación:**

**Objetivo general:**

Analizar la actitud de los directivos nacionales de la industria manufacturera respecto de la innovación y creatividad en sus organizaciones.

La redacción de los objetivos debe encaminar la investigación a lo que se desea alcanzar.

**Objetivos específicos:**

1. Elaborar un marco teórico sobre la actitud de los directivos de empresa respecto de la innovación y creatividad en las organizaciones productivas.
2. Diseñar un instrumento de medición de aptitudes con sus respectivas pruebas de validez y confiabilidad para aplicarlo a la población seleccionada.
3. Identificar la actitud de los gerentes de las 100 empresas más grandes del país respecto de la innovación y la creatividad en ellas.
4. Analizar los resultados obtenidos según subsector y demás variables sociodemográficas de las empresas estudiadas.

Los objetivos van relacionados con verbos en infinitivos que describan la acción que se desea llevar a cabo en el proceso de investigación.

Los objetivos están estrechamente ligados al problema de investigación.



## Elementos constitutivos de la metodología de la investigación actual

Elementos constitutivos de la metodología de la investigación actual. En la metodología de la investigación actual se utilizan sus estructuración elementos como las variables, los constructos, los indicadores de ellas y los índices. Esto implica que los análisis que se hagan de los datos procesados durante de la investigación ejecutada se harán sobre elementos. Entones haremos una introducción a cada uno de ellos.

### Variables

Las variables son características o propiedades de la realidad que pueden fluctuar dentro de unos valores en individuos, conjuntos o poblaciones. El peso, la edad, la talla, la ocupación son consideradas variables debido a que, en una muestra o conjunto de individuos, dichas propiedades pueden presentarse en grados diferenciados entre dichos integrantes.

### Clasificación de variables

#### Variables dependientes, independientes e intervinientes

La clasificación más extendida y general de las variables las dividen en dependientes e independientes. En los proyectos de investigación suelen encontrarse las denominadas variables intervinientes. Estos tres tipos de variables se definirán a continuación.

1. Variables Dependientes: son las variables que muestran las consecuencias o los efectos que producen los cambios de valor o modalidad de las variables independientes sobre un problema o fenómeno.

2. Variables Independientes: son aquellas características o propiedades que originan un efecto en otra variable relacionada, por tanto, se identifican como las causantes de un problema o fenómeno. Las variables independientes son las denominadas variables experimentales dentro de un proyecto de investigación. Por ejemplo, si se busca estudiar la incidencia de cáncer de estómago (variable dependiente) en personas que consumen alimentos transgénicos (variable independiente).

Las variables pueden ser:

- Variables Dependientes
- Variables Independientes.
- Variables Intervinientes.
- Variables Cuantitativas
- Variables Cualitativas.
- Variables individuales.
- Variables Colectivas.
- Variables nominales.
- Variables Ordinales.
- Variables Proporcionales.

Variables Intervinientes: son aquellas variables que funcionan como modificadores positivos o negativos entre las variables independientes y las variables dependientes. Como las variables intervinientes no son medibles o manipulables, en los estudios de investigación es necesario mediante la observación identificarlas para evitar atribuir solo a las variables independientes el efecto causado por la combinación de ambas. Ejemplos de variables intervinientes pueden ser el entusiasmo o motivación de un empleado para ejecutar una actividad. La experticia (variable independiente) de este trabajador a ejecutar una actividad crea un producto (variable dependiente) El entusiasmo o la motivación (variables intervinientes) pueden influir en una

mayor o menor calidad del producto terminado.

### **Variables cualitativas y cuantitativas**

4. Variables cualitativas: son aquellas propiedades de los objetos o elementos de estudio que no pueden ser medibles en términos de cantidades sino como características que se encuentran presente o no en ellos. Entre ellos pueden estar la religión, el sexo, actividad económica, idioma, etc.

5. Variables cuantitativas: son aquellas variables que son susceptibles a ser medidas en términos numéricos, como la edad, la estatura, el peso, el salario, etc. Pueden presentarse de dos maneras: continuas o discontinuas.

5.1. Continuas: son aquellas que pueden tomar un valor cualquiera dentro un rango definido.

5.2. Discontinuas o discretas: son aquellas que solo toman valores enteros o un cierto número de valores finitos debido a que la unidad de medición no puede ser fraccionada. Ejemplo numero de estudiantes en un salón de clases.

### **Variables individuales y colectivas**

La siguiente clasificación de variables tiene una importante presencia en el desarrollo de investigaciones sociales cuantitativas, estas son las variables individuales y las variables colectivas.

6. Variables individuales: son propiedades presentes en los individuos, como por ejemplo su dominio de idiomas.

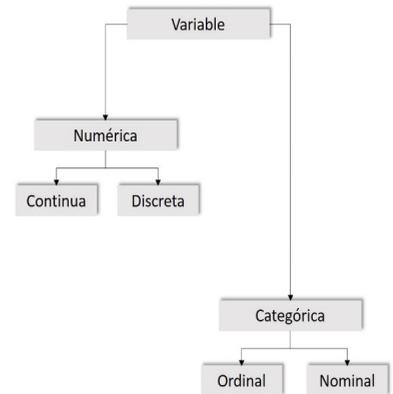
7. Variables colectivas: son aquellas propiedades presentes en grupos de individuos o segmentos sociales que se fundamentan en las propiedades presentes en los integrantes de ese grupo o segmento social. La representación de dichas propiedades se alcanza a través de la aplicación de operaciones aritméticas sobre el segmento o grupo social. Siguiendo el ejemplo del dominio de idiomas la tasa de dominio de idiomas de segmento o grupo social será el promedio de idiomas que dominen sus integrantes. El nivel de emprendimiento del grupo o segmento será en porcentaje de emprendimientos que ejecuten sus integrantes.

Es importante destacar que las técnicas utilizadas para analizar las variables individuales son susceptibles de ser utilizadas también con las variables colectivas. Esto se hará más evidente a medida que se avance en el estudio de estas variables.

### **Variables nominales, ordinales y proporcionales**

Estas variables se apoyan en las propiedades matemáticas que poseen los números y la manera en que estadísticamente pueden ser manipuladas a fin de asignar grados o modalidades a las características que pueden manifestar los individuos y los colectivos. Entre ellas tenemos las variables nominales, ordinales y proporcionales.

Variables nominales: en estas variables manipulan propiedades cualitativas como sexo, estado civil, etc. En las variables nominales, el uso de nomenclaturas numéricas sirve únicamente para diferenciar unos individuos de otros; por ejemplo, clasificar con el valor “1” a las mujeres solteras y “2” a los hombres solteros solo realiza un papel identificador, por lo que cambiar la nomenclatura no incide en los resultados de la investigación y por tanto solo sirve para conocer cuántos individuos están en un momento determinado en alguna de las dos categorizaciones.



9. Variables ordinales: en las variables ordinales las propiedades o características atribuibles a un grupo de individuos pueden tomar diferentes valores dentro de escala jerárquica, esto quiere decir que hay relaciones de mayor que y de menor que entre los integrantes del grupo objetivo. Ejemplo se puede abordar 4 niveles de salarios, 3 niveles de participación comunitaria, etc. La asignación de valores numéricos a las jerarquías solo indica que un individuo puede poseer más poder adquisitivo que otro sin conocer magnitud de la diferencia.

Variables proporcionales: estas variables se distinguen por su nivel de medición, en estas los valores asignados a las distintas jerarquías poseen propiedades aritméticas, por tanto, existe un valor mínimo, que puede ser cero hasta un valor máximo.

Dentro de este intervalo se cuantifica el nivel de presencia de una característica del individuo o grupo objeto de estudio, mediante su proporción. Si el salario de un individuo es 500 en un momento determinado y el salario de otro individuo en el mismo momento de tomar el valor anterior es de 250, entonces, el primer individuo posee el doble de recursos que el segundo, o que el segundo posee la mitad de recursos que el primero. También se pueden promediar los salarios, sumarlos, etc.

### Elaboración de constructos

Los constructos son aquellas cualidades que posee una persona y que puede ser objeto de estudio. En este sentido, Gras afirma que cuando una cualidad (inteligencia, agresión, tendencia, rendimiento, creatividad, etc.) puede ser medida, observada y además relacionada con otros conceptos a través de un enunciado o hipótesis, puede ser utilizada para llevar adelante un proyecto de investigación científica. En esto radica la importancia del constructo. Los constructos están constituidos por propiedades que no son evidentes y no se pueden medir de forma directa, sino que se miden a través de su manifestación.

### Indicadores

Hay que notar que, conociendo la definición de variable, se puede concluir que los constructos son también variables; sin embargo, como están formados por variables subyacentes que no son medibles de forma directa requieren de un sistema que permita ponderar las manifestaciones externas de dichas variables. A este sistema se le denomina indicadores.



En el caso de una cualidad, los indicadores de ella se construyen a partir de preguntas que se deben asociar con las propiedades de las variables subyacentes que le definen; las respuestas a una o más preguntas son las que permitirán evaluar o valorar (cuantificar) la cualidad estudiada.

### **Operacionalización de los indicadores**

Cuando se identifican los indicadores específicos que cuantificaran una variable subyacente, se está operacionalizando dicha variable. Esto quiere decir que se va a definir un concepto abstracto o teórico mediante indicadores operacionales, quienes pueden ser verificables y medibles. Los indicadores se operacionalizan a través de los instrumentos, que son sistemas que permiten realizar mediciones sobre cada indicador. Sobre estos se puede decir que para construirlos el proceso que se debe seguir es dividir las variables subyacentes a otras variables más específicas susceptibles a ser medidas, estas variables son denominadas dimensiones o componentes; la característica principal de una dimensión es poder ser medido, cuantificable por lo que puede ser observada de manera directa.

Para clarificar, Por ejemplo, si el concepto que se quiere medir es la coordinación corporal necesaria para manejar, entonces se requiere verificar varias variables como el estado de la vista, los reflejos, la coordinación brazo-pierna, etc. Para ello el instrumento utilizado es el test psicosenométrico.

En los párrafos siguientes expondremos algunos ejemplos de operacionalización sobre variables subyacentes.

a) Si se quiere conocer la definición operacional de la satisfacción del cliente, se le asigna una valoración a ciertas preguntas que dimensionen dicha satisfacción. Una manera sería como se muestra a continuación.

- El cliente siempre consigue los productos que necesita:

Siempre: 2. Casi siempre: 1. Nunca: 0.

- El cliente ante dudas sobre un producto, consigue asesoría adecuada:

Siempre: 2. Casi siempre: 1. Nunca: 0.

- El cliente recibe el catálogo de productos en oferta con frecuencia:

Siempre: 2. Casi siempre: 1. Nunca: 0.

En este ejemplo si un cliente suma 6 puntos, es un cliente que será poco propenso a cambiar de establecimiento, ya que se siente cómodo con la atención que le brindan. Si otro cliente suma 2 o menos, es un cliente que potencialmente puede abandonar al establecimiento ya que no se siente conforme con el servicio que le brindan. A escalas mucho más grandes, estos datos pueden ser cruciales para las perspectivas de crecimiento y sobrevivencia del establecimiento y podría regir las medidas a tomar para atraer más clientes o para no perderlos.

b)La definición operacional de nivel social familiar. Se redactan las preguntas que pueden dimensionar



esta variable.

•Grado académico alcanzado:

Universitaria: 3 Secundaria: 2. Primaria: 1

•Poder adquisitivo mensual

Alto: 3 intermedio: 2 Bajo: 1

•Número de integrantes de la familia:

Dos integrantes: 4 de tres a cuatro integrantes: 2 más de 4 integrantes: 1

Se puede observar que, dependiendo de las combinaciones que pueden arrojar los datos, las personas que a pesar de poseer un título universitario tienen una familia numerosa y un ingreso medio, tendría un nivel social familiar inferior a una persona con un nivel educativo de secundaria, ingresos medios y una familia pequeña de menos de tres o menos personas.

## Índices

Los índices son medidas de variables complejas que son producto de la combinación de dos o más indicadores o variables. El investigador es quien escoge mediante un proceso en el cual se identifican aquellas características que mejor describen a una variable compleja, los indicadores o variables adecuadas para valorarla. Entonces, si se quiere conocer el índice, por ejemplo, de la satisfacción del cliente, lo más adecuado sería establecerlo a partir de la combinación de las valoraciones que se obtienen de los indicadores de satisfacción sobre productos preferidos y satisfacción sobre ofertas de productos.

## Hipótesis

En muchos proyectos de investigación bajo la metodología cuantitativa, el objetivo o los objetivos de investigación se pueden plantear como una o más hipótesis. Una hipótesis es una suposición sobre qué cualidades debería tener en la realidad la respuesta a la pregunta principal de investigación. En las ciencias sociales una hipótesis le proporciona a la problemática estudiada una perspectiva acerca de cómo debería ser la respuesta o comportamiento del fenómeno social en el entorno en el cual se desarrolla,

En las ciencias sociales una hipótesis le proporciona a la problemática estudiada una perspectiva acerca de cómo debería ser la respuesta o comportamiento del fenómeno social en el entorno en el cual se desarrolla, qué se puede esperar del mismo según lo que se podría esperar de las respuestas de las variables que la conforman

Lo indicado antes es muy importante subrayarlo, la hipótesis nace dentro de un marco de conocimientos que se han verificado sobre aspectos de la problemática que se está estudiando; por tanto, toda conjetura que no esté dentro de dicho marco de conocimientos no puede llamarse hipótesis.

La importancia de la formulación de una hipótesis estriba en la capacidad de la misma de servir de enlace entre el conocimiento teórico que se tiene sobre el tema y la investigación empírica que permite la creación de nuevo conocimientos y con ello, nuevos aportes a un área de la ciencia. Una hipótesis sirve para orientar los esfuerzos hacia ciertas áreas que podrían brindar los resultados esperados en detrimento de otros, consiguiendo ciertos patrones de regularidad en los hechos estudiados. En última instancia podría dar respuesta sobre la problemática abordada en el proyecto de investigación.

Las hipótesis nacen de conjeturas que el investigador o los investigadores plantean a partir de estudios previos entorno a la problemática estudiada, o de suposiciones o conjeturas de los propios investigadores como producto de un análisis personal sobre teorías que, bajo ciertas condiciones produjeran ciertos resultados.

Una formulación de una hipótesis podría ser de esta manera.

Ejemplo de un objetivo formulado como hipótesis

Objetivo: En la provincia de Arequipa, encontrar la relación existente en el número de microempresas existentes con los niveles de oferta de empleo en la región.

Hipótesis: En la provincia de Arequipa, el nivel de emprendimiento a nivel de microempresas está directamente relacionado con la baja calidad de la oferta de empleo ofrecido por medianas y grandes empresas.

### **Clasificación de las hipótesis**

La hipótesis se puede plantear según las características de la investigación que se desea desarrollar, por ello y para una simplificación en su clasificación, se asocian los criterios y relaciones que debe haber entre la hipótesis con la pregunta de investigación y con el desarrollo de objetivos sobre dicha pregunta. Teniendo en cuenta esta relación, se pueden clasificar las hipótesis como se muestra a continuación.

### **Hipótesis descriptivas**

Este tipo de hipótesis intentan realizar una predicción sobre la posibilidad de que ciertos sucesos se estén produciendo y no están siendo detectados o que posiblemente ocurran en el futuro. Por ejemplo, si una persona no es fumadora a la edad de 18 años, su posibilidad de serlo es menor al 20%.

La hipótesis es la suposición que permite establecer relaciones entre hechos y, de esta manera explicar por qué se produce.



## **Hipótesis generales o teóricas**

Son aquellas hipótesis que aprovechan el diseño conceptual que los científicos han construido sobre algún fenómeno que desean investigar. Estas hipótesis nacen de los procesos de observación y de las generalizaciones hechas sobre algún sujeto de estudio que usualmente anteceden a los procesos de investigación que se podrán desarrollar posteriormente, lo que implica que no se han cuantificado propiamente todas las variables que podrían estar involucradas. Ejemplo, las personas con dominio de dos a más idiomas tienden a ganar más que los que dominan un solo idioma.

## **Hipótesis tipológicas**

Son aquellas que proponen clasificaciones sobre el objeto de estudio que abordan en la investigación. Ejemplo, con respecto a la diversificación de las entregas a domicilio, las aplicaciones pueden clasificarse en:

- a)Bajo demanda
- b)Corporativas
- c)Focalizadas según grupos de interés
- d)A medida

## **Hipótesis relacionales**

En estas hipótesis se destaca las relaciones que existe entre dos o más variables presentes en la problemática u objeto de estudio. “El nivel educativo tiene relación inversa con la ingesta de grasas saturadas y directa con el consumo de cereales y vegetales”

Dentro de esta categoría están las hipótesis complejas, que relacionan diferentes variables del objeto de estudio según ciertos índices o indicadores. La altura es producto de la combinación de la alimentación, accesos a salud y cultura física.

## **Hipótesis direccionales y no direccionales**

Las hipótesis direccionales indican cómo debería ser la respuesta esperada de un objeto de estudio según las características del mismo. Por ejemplo, Los internautas prefieren declinar el acceso a sus datos personales cuando los sitios web muestran abiertamente sus políticas de uso de datos.

En las hipótesis no direccionales los investigadores reconocen las relaciones existentes entre variables que definen al objeto de estudio, pero no estipulan la gradualidad de dichas relaciones. Por ejemplo, existe una respuesta inmunológica positiva de un neonato cuando entra en contacto directo con la piel de su madre

## **Hipótesis explicativas o causales**

Estas se orientan a la explicación de ciertos acontecimientos que se producen o pueden producirse alrededor o, al propio objeto de estudio, enfocándose en la determinación del por qué de su incidencia o posible incidencia, ubicando las relaciones causas y efectos que las determinan. “La incidencia del alcoholismo en jóvenes es mayor cuando sus padres o familiares son consumidores habituales”.

Es común en muchas hipótesis explicativas que se proponga una sola fuente causal como detonante de una problemática; sin embargo, la experiencia social confirma que son múltiples los factores que inciden en la manifestación de una problemática particular. Por ello es necesario acotar las investigaciones de este tipo y normalizar las condiciones para enfocarse en la incidencia de algunas de las variables involucradas. Por ejemplo, si se quiere estudiar la religiosidad de una población, se debe acotar la libertad de culto, suponiendo igualdad de libertad para todos los individuos de escoger de manera libre.

## **Hipótesis estadísticas**

Son aquellas que transforman en valores estadísticos mensurables, con sus respectivos márgenes de error, una hipótesis de investigación que utiliza muestras de una determinada población. Las hipótesis estadísticas poseen dos baremos de medida; una, denominada la hipótesis nula, la que indica que dentro de una población no existe una variación importante en el promedio, como, por ejemplo, que, entre actores de una serie de éxito, no hay variaciones importantes en los salarios entre los protagonistas masculinos y los femeninos, y otra, denominada la hipótesis alternativa, que afirma que, entre los protagonistas de una serie de éxito, sí existen diferencias en sus salarios. Para convalidar si una u otra alternativa es la que se presenta en el estudio, se deberá conocer la probabilidad de que un caso u otro sea el más factible de producirse. Para ello es el investigador el que propone cómo medir los valores para conocer el resultado.

Dentro de las hipótesis estadísticas están las hipótesis empíricas, que son aquellas que se caracterizan por obtener resultados que pueden comprobarse o descartarse de forma estadística. Por ejemplo, los ingenieros matemático-computacionales tienen un índice de desocupación laboral menor a 1%. Hay muchos chicos pobres y marxistas en mi universidad.

## **Características que debe tener una hipótesis**

-Deben existir relaciones entre dos más variables, existiendo variables independientes y variables dependientes.

-Las relaciones deben tener la capacidad de poder ser puestas a prueba, poder ser medibles y verificables y por tanto capaces de ser observadas.

-Las hipótesis deben poder realizar predicciones sobre el comportamiento de un objeto de estudio y explicar dicho comportamiento.

-Los resultados que se desprenden del planteamiento de la hipótesis deben ser justificables, coherentes con teorías y experiencias anteriores, correspondiendo con una línea de estudio sistemático.

-Las ideas que dan sustento a una hipótesis deben ser claras, comprensibles y cuya interpretación sea



siempre la misma independiente del individuo que la lea.

- Deben redactarse de forma afirmativa, evitando expresiones que propongan juicios de valor o sesgos.
- Deben ser concretas, especificando con claridad las relaciones entre variables.



# CAPITULO 3

## La investigación cuantitativa y sus diseños

Para responder a las preguntas de investigación se requiere diseñar un proyecto que se construya alrededor del objeto investigado. En este sentido, la bibliografía disponible para la construcción de proyectos con enfoque cuantitativo indica que existen tres caminos o tipos concretos de investigación que se enfocan en la posibilidad que tiene el investigador de manipular bajo cierto control, la o las variables independientes, entre otros factores (como podría ser el control sobre la población objeto de estudio). Estos caminos son los siguientes.

- a. Investigación cuantitativa experimental
- b. Investigación cuantitativa cuasi-experimental
- c. Investigación cuantitativa no experimental

Dentro de esta clasificación se pueden encontrar subgrupos de investigaciones cuantitativas como pueden ser las investigaciones descriptivas o las investigaciones explicativas que se encuentran en el grupo de investigaciones experimentales y se diferencian según sea el objetivo principal de la investigación. Si se toma el tiempo de duración podríamos tener investigaciones de tipo sincronicas o transversales, que son las que toman muestras en un único momento, o las investigaciones longitudinales, en donde a la población que es objeto de estudio se le toman datos durante un intervalo de tiempo continuo. Estas investigaciones forman parte de las investigaciones no experimentales.

### **Investigación cuantitativa experimental**

Este tipo de investigación busca determinar con el mayor grado de seguridad posible las relaciones existentes entre las causas y las consecuencias observadas en una problemática objeto de estudio. Los investigadores poseen un grado de control sobre las variables independientes estimuladoras detectadas en la problemática de estudio, que pueden hacer fluctuar según los requerimientos que la investigación necesite según los objetivos buscados.

En esta línea, los investigadores conforma varios grupos de estudio que se pueden categorizar como experimentales y de control. A quienes se les suministrará variaciones de valores de entrada en las variables independientes son los grupos experimentales; los resultados obtenidos en dichos grupos, que fueron sometidos a las variaciones planificadas por los investigadores en las variables independientes, se contrastan con otros grupos que no han recibido ningún estímulo; estos últimos son denominados grupos de control.

Ejemplos de investigaciones cuantitativas experimentales

Pueden servir de ejemplo de investigaciones experimentales las siguientes propuestas

- Verificación de la efectividad de la cloroquina como tratamiento contra el paludismo.



- Investigar los efectos de la radiación de Cloruro de Cesio 137 sobre el agua.
- Probar que el herbicida RoundUp hecho a base de glifosato es cancerígeno.

### **Alcances de las investigaciones experimentales**

Las investigaciones experimentales en su mayoría permiten establecer:

a. Las consecuencias que tiene la variable independiente sobre la dependiente. Ejemplo: la influencia del nivel de dislexia en niños mayores a 12 años sobre el aprendizaje de idiomas.

Las consecuencias de la combinación de variables independiente sobre otras variables. Por ejemplo: efecto de la curva de aprendizaje en manejo vehicular cuando se utilizan simuladores de conducción y prácticas reales.

### **Características de la investigación experimental**

• Las investigaciones experimentales requieren mucha rigurosidad al momento de manipular las variables o factores propios del experimento y de controlar aquellos aspectos o factores no relacionados al experimento que pueden incidir en el correcto desarrollo del mismo. Para poder ejercer control sobre los aspectos o factores mencionados anteriormente se utiliza el control directo, eliminando los factores de contaminación de la muestra, o por procedimientos estadísticos aleatorios que incluyen la selección aleatoria de los individuos que compondrán la muestra o población objeto de estudio; aleatoriedad en la conformación de los grupos de control y los experimentales y por último, asignación aleatoria del tratamiento experimental a cualquiera de los grupos conformados.

b. Las consecuencias que tienen dos o más variables independiente comparadas diferencialmente sobre una variable dependiente. Ejemplo: comparación de la curva de aprendizaje en el manejo vehicular cuando se realizan prácticas de manejo reales con las curvas de aprendizaje obtenidas cuando se realizan prácticas de manejo simuladas.

En la investigación experimental existe un conjunto de variables que se dejan fijas o constantes, para contrastar con otras que se consideran el sujeto de la investigación.

Las investigaciones experimentales se aseguran que tanto los grupos experimentales como los grupos de control posean características homogéneas. Esto con el fin de no introducir factores que generen distorsiones que afecten el experimento. La única diferencia estribará en que a los grupos experimentales se les aplicarán los factores causales (introducción de las variables independientes a los grupos experimentales) y a los grupos de control no.

- La investigación experimental es la más recomendada para investigar aquellas relaciones que presentan un conjunto de consecuencias producto de un número limitado de causas. Sin embargo, el limitar o acotar los factores causales para el diseño de un experimento produce un efecto restrictivo en los resultados de investigación, lo que produce una sensación de artificialidad en ello, que es lo mismo decir que no reproduce la realidad que los seres humanos experimentan, bien sea por las condiciones de observación que requieren ambientes controlados o por razones éticas que interponen una cierta moralidad al desarrollo de una investigación.

### **Etapas de una investigación cuantitativa Experimental**

Una investigación cuantitativa por lo general posee el siguiente esquema de etapas en su proceso de desarrollo e implementación.

1)Se consulta la literatura que corresponde con la línea de investigación en la cual se encuentra la problemática que se quiere abordar. Sirve como paso necesario para acortar el alcance que tendrá la investigación.

2)Una vez acotada la problemática, se realiza la identificación y definición del problema.

3)Se formula la hipótesis entorno a la cual girará la investigación, deduciendo los resultados que podrían ser medibles u observables, definiendo los términos básicos para ello.

4)Se elabora el proyecto de investigación, el cual implica:

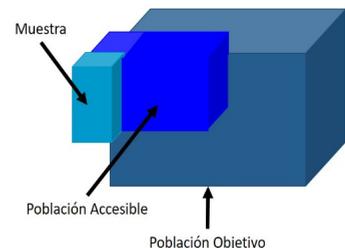
Identificar los factores o variables no experimentales que podrían afectar el experimento y definir la manera de controlarlos.

- Escoger el diseño experimental adecuado. Este apartado sera tratado en una sección posterior.

- Definir el tamaño de la muestra adecuada para modelar apropiadamente al objeto de estudio, asignarlos a los grupos tanto de control como experimentales. Luego se aplica el tratamiento experimental (manipulación de variables independientes) al grupo experimental.

- Se construyen o se escogen los instrumentos que serán utilizados para realizar el experimento y luego pasar a la medición de los resultados.

- Se construyen los procedimientos que permitirán la recolec-



ción de los resultados del experimento. Dichos resultados son los datos.

- Se enuncia la hipótesis nula. Que es aquella que sirve de contraste con la hipótesis experimental.

5)Se procede a realizar el experimento.

6)Se organizan los resultados de forma estadística, a fin de reconocer el patrón de referencia que demuestre el efecto o las consecuencias del experimento.

7)Aplicar las pruebas estadísticas apropiadas para interpretar los resultados.

8)Realizar un informe escrito donde consten los resultados obtenidos en el proyecto, para su respectiva socialización.

### **Diseño experimental en las investigaciones cuantitativas**

En este apartado se presenta las formas más habituales de encarar el proceso de experimentación en una investigación cuantitativa

A)Cuando el experimento se diseña con dos grupos conformados de forma aleatoria, uno de control y otro experimental, donde se realizan mediciones tanto antes como después de aplicar el estímulo.

Este es el diseño más habitual en las investigaciones cuantitativas. En él se asignan elementos a los grupos experimental y de control de forma aleatoria, sin seguir ningún patrón preestablecido. Una vez asignados todos los elementos de los grupos se prosigue de la siguiente manera:

a.1) Como se indica en el enunciado, se realizan mediciones en ambos grupos, antes de introducir algún estímulo al diseño. El objetivo es conocer, en ambos grupos, las condiciones iniciales de la variable dependiente, quien es la que refleja el efecto del estímulo que introduce la variable independiente o experimental.

a.2) Se aplica el estímulo de la variable independiente solo en el grupo experimental. (ejemplo: al grupo de control se le da un medicamento mientras al otro solo un placebo).

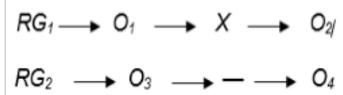
a.3) Se realizan mediciones en los dos grupos y se anotan los resultados.

a.4) Se realizan las comparaciones entre las mediciones recopiladas antes del estímulo y después del estímulo en cada grupo.

El estímulo aplicado por la variable independiente puede tener dos modalidades, de corta y de mediana duración. El estímulo de corta duración puede ser como, por ejemplo, un test de aerodinámica sobre un automóvil en un túnel de viento, una clase magistral, etc. El estímulo de mediana duración pudiese prolongarse por unas semanas como podría ser, por ejemplo, un circuito de entrenamiento para mujeres en la etapa de posparto. En este último caso, se deben controlar las influencias externas que puedan alterar los resultados, invalidando así los datos recopilados en el grupo experimental dado que no se tendría la certeza de que las variaciones observadas se deban a los cambios de la variable experimental de forma directa durante el experimento.

Entonces, las relaciones entre los elementos antes expuestos se pueden escribir de la siguiente manera

Para tener la confianza de que los resultados de la variable dependiente son causados directamente por la variable independiente o experimental, existe el concepto de validez interna, que no es otra cosa que verificar que lo antes expuesto es real. Donald Campbell y Julian Stanley en su libro Diseños experimentales y cuasi-experimentales en la investigación social, proponen una metodología para analizar este aspecto. La simbología que se utilizará para ejemplificar el experimento estándar se detalla a continuación:



R: asignación de individuos a un grupo de forma aleatoria (random).

E: emparejamiento o nivelación.

G: grupo de individuos.

X: aplicación del estímulo.

O: medición de los individuos de un grupo.

—: ausencia del estímulo en la variable independiente (grupo de control).

Al conformar grupos de forma aleatoria se elimina cualquier parcialidad que pudiese afectar las mediciones; luego se les aplica una prueba a ambos grupos antes y después de introducir un estímulo en el grupo experimental y ninguno en el grupo de control, y si existen diferencias al comparar los resultados, y si están se inclinan a favor del grupo experimental, se puede deducir que los cambios son producto de la influencia del estímulo aplicado por la variable independiente a dicho grupo.

Debido al uso de la aleatoriedad en la conformación de los grupos, se controlan los factores de parcialidad o intencionalidad en los grupos; a esto se le denomina control de factores de selección. La selección aleatoria también controla el factor de mortalidad, que se refiere a que, si dos grupos están constituidos de forma similar, el grado de deserción de individuos que no llegaron a completar el experimento será muy pa-

recido en ambos grupos. En esta misma línea pensamiento, la selección aleatoria también permite controlar los factores asociados con la historia, propia de cada individuo, los cuales son acontecimientos externos al experimento y que podrían impactar sobre los resultados. La aleatoriedad permite una distribución similar a cada grupo, controlando su efecto. También se controlan los efectos de maduración, que son aquellos cambios biológicos y psicológicos en los individuos que podrán producirse durante el experimento, que por la selección aleatoria permite que se distribuyan homogéneamente en cada uno y no ser un agente de distorsión.

Para conocer si existen diferencias entre los grupos de control y experimental después de realizado el experimento, es necesario aplicar primero un análisis de covarianza sobre los valores que se conocían antes de la aplicación del estímulo como los elementos de la covariable. Con ello se logra filtrar las diferencias en los individuos de cada uno de los grupos, permitiendo que solo los cambios atribuibles a la variable independiente queden como causantes de las variaciones que se puedan observar después de la aplicación del estímulo en el experimento. Sin embargo, si existe la certeza de que las condiciones de ambos grupos al momento de iniciar la experimentación poseen valores muy similares en lo que respecta a los valores de la variable dependiente, entonces se ejecuta un análisis estadístico “t” o de varianza directamente.

B) Cuando el diseño experimental se conforma con un grupo de control, un grupo experimental y con mediciones únicamente después de aplicado el estímulo.

En este diseño no se realizan mediciones antes de la aplicación del estímulo a la variable independiente. Las razones para ello pueden encontrarse en una incapacidad real o inducida para determinar en cada grupo los valores de las variables antes de realizar el experimento. En este tipo de diseño se cumplen las siguientes etapas.

b.1) Se sigue la pauta de realizar una selección de individuos para cada grupo, de control y experimental de manera aleatoria. Según las exigencias del proyecto de investigación, puede existir más de un grupo experimental (ejemplo, dos sistemas de ejercicios aeróbicos para mujeres en etapa posparto).

b.2) Se aplica el estímulo a la variable independiente únicamente sobre el(los) grupo(s) experimental(es).

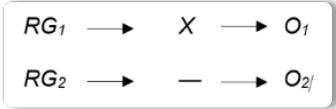
b.3) Se realiza la medición de efecto sobre la variable dependiente después de haber sido aplicado el estímulo (ejemplo, estado de ánimo de las mujeres luego de realizado los ejercicios aeróbicos) tanto en el grupo de control como en el(los) grupo(s) experimental(es).

b.4) Comparación de los resultados obtenidos después de aplicado el estímulo en ambos grupos.

Como en el diseño clásico, en este diseño se puede controlar factores de perturbación de datos como lo pueden ser la historia, la maduración, la selección y la mortalidad, Estos factores que atentan contra la validez interna del experimento pueden ser minimizados debido al empleo de un grupo de control seleccionado de forma aleatoria; sin embargo y debido a que no se realizan mediciones en los grupos de control y experimental antes de la aplicación del estímulo a la variable independiente del diseño, no se puede con certeza el efecto del mismo. Entonces el diseño se encuentra con una serie de desventajas como

El diseño del experimento se puede expresar simbólicamente como sigue.

b.4.1) No se puede realizar un análisis de covarianza para mitigar las diferencias iniciales entre el grupo de control y el(los) grupo(s) experimental(es).



b.4.2) Si se produce una variación en la cantidad de individuos en cada grupo, y al no haber mediciones iniciales antes de introducir el estímulo, no se puede conocer si los efectos medidos después de aplicado el estímulo responden únicamente a él o si parte de la respuesta obedece a las diferencias existentes antes entre los grupos de control y experimental (les).

El análisis del experimento se ejecuta sobre los datos obtenidos después de aplicado el estímulo a la variable independiente, observando sus efectos sobre el(los) grupo(s) experimental(es), para ello se utiliza la estadística “t” si es un solo grupo experimental; si son dos o más grupos se realiza un análisis de varianza.

C)Diseño experimental mediante el emparejamiento directo entre los grupos de control y experimental y mediciones antes y después de aplicar el estímulo.

Similar al diseño experimental tradicional, este diseño varía en la construcción de los grupos de control y experimental. Si en el diseño tradicional la selección se realizaba mediante la asignación aleatoria de individuos a cada grupo, en este diseño la selección de los individuos de cada grupo se ejecuta mediante emparejamiento de pares de individuos, cuya correspondencia directa busca, de manera similar al diseño tradicional, eliminar en lo posible las diferencias entre los grupos de control y experimental. La metodología que sigue este diseño es el siguiente.

c.1) Se realiza un test exploratorio a un grupo aleatorio de individuos destinado a establecer una relación entre su respuesta (la construcción del test busca medir una variable dependiente -por ejemplo, su posición acerca de la religión- relacionada directa o de forma estrecha con el estudio que se desea ejecutar) y la variable independiente que actúa de estímulo.

c.2) Una vez aplicado el test y obtenido los resultados por cada individuo, se procede a emparejar aquellos individuos cuyas reacciones resultaron iguales o muy similares, donde las diferencias resulten descartables.

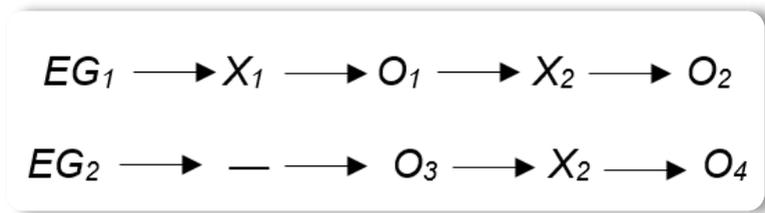
c.3) Realizado el emparejamiento se asigna uno al grupo de control y el otro al grupo experimental.

c.4) Una vez conformados los grupos, se aplica al grupo experimental el estímulo planificado (por ejemplo, un video sobre el origen de diversas religiones), y al grupo de control ningún estímulo o un estímulo tendiente a nulo (como por ejemplo un placebo en un estudio médico).

c.5) Para una medición posterior a la aplicación del estímulo se realiza el test inicial -el aplicado cuando los grupos aún no estaban conformados- tanto en el grupo de control como en el experimental, o uno similar, y se registran los resultados.

c.6) Se procede luego a la comparación de los resultados y se establecen las diferencias obtenidas después de aplicado el test tanto en el grupo de control como en el experimental.

Se puede representar el diseño experimental de forma simbólica de la siguiente manera



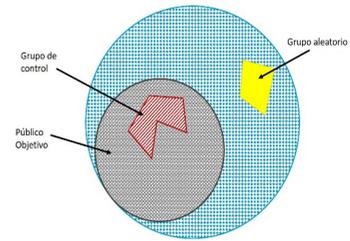
Donde X2 es el test preselección de grupos o uno similar.

Para la ejecución del análisis diferencial de los resultados se utiliza el modelo estadístico “t” la cual es la relación que existe entre la desviación del valor estimado con respecto a su valor hipotético y su error estándar.

Es importante resaltar que en general se han presentado los diseños experimentales con una sola variable independiente; sin embargo, se debe considerar que, en muchas investigaciones, y sobre todo en el campo de la investigación educativa y de las líneas de enseñanza-aprendizaje, existen investigaciones que pueden utilizar dos o más variables independientes para evaluar los efectos combinados de ellas sobre la variable dependiente que es motivo del estudio. A este tipo de diseño experimental se le denomina diseño

factorial. Como ejemplo existen los diseños factoriales  $2 \times 2$ ,  $m \times n$  y el denominado diseño jerárquico.

Otro aspecto importante es el concerniente al tamaño de los grupos tanto de control como experimentales. El tamaño de los grupos se decide teniendo en cuenta el modelo estadístico que se utilizará en el análisis de datos, la estimación de la varianza, la significación establecida y el error estándar presente en la tabulación de datos y quien reflejará la diferencia real entre los valores experimentales y los hipotéticos. Sin embargo, y sin detrimento de estos criterios previos, los investigadores se decantan por grupos de 10 individuos, variando entre 5 y 30 individuos.



### Investigaciones cuasiexperimentales

En la gran mayoría de diseños experimentales, los grupos de control y de experimentación son escogidos mediante una designación aleatoria de los individuos que los compondrán, que luego de realizada la selección, ambos poseerán un número de integrantes iguales y características similares, lo que indica que ambos grupos son equivalentes. La asignación aleatoria busca, como ya se ha mencionado, minimizar la influencia de dichas características propias de los individuos, que podrían añadirse a la variable independiente y que ocasionarían alteraciones en las mediciones a realizar sobre la variable

En la población encontraremos el público objetivo, los grupos aleatorios y de control, necesarios para el estudio

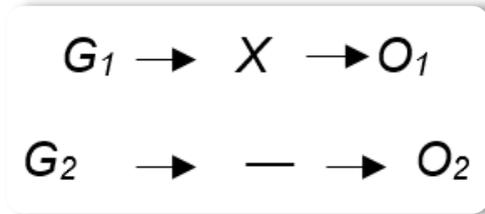
En la realidad se presentan con frecuencia, circunstancias que impiden que el proceso de selección aleatoria de los individuos, que luego pertenecerán a los distintos grupos, se ejecute satisfactoriamente, lo que dificulta el proceso de control y anulación de características que todo diseño de investigación requiere. Sin embargo, la literatura existente sobre los diseños de investigación nos señala que, a falta de la posibilidad de seleccionar aleatoriamente los individuos de los grupos de control y experimental, aún existen posibilidades de diseñar un experimento que a pesar de la limitación mencionada sigan teniendo tanto coherencia y validez interna como externa. Estos son los diseños de investigación cuasiexperimentales.

Los diseños de investigación cuasiexperimental poseen elementos similares a los diseños de investigación experimentales; sin embargo, las comparaciones de las respuestas obtenidas de los individuos se ejecutan entre grupos que no son equivalentes o grupos intactos, lo que quiere decir que no se escogen al azar y por ello pueden tener diferencias notables entre el grupo experimental y el de control, aún antes de ser expuesto el grupo experimental a un estímulo. Adicional a esta situación es probable que incluso no exista un grupo de control como tal, lo que añade otro nivel de dificultad al intentar conocer el efecto de la variable independiente sobre la variable dependiente que se pretende medir.

Conociendo lo anterior, los diseños cuasi experimentales tienen en general dos modelos de ejecución, que se muestran a continuación

#### A)Diseño cuasiexperimental con un grupo de control no equivalente

Este diseño está formado por dos grupos, donde el grupo seleccionado como experimental, como por ejemplo los estudiantes de la escuela de enfermería de una universidad, es quien recibirá el estímulo experimental, el grupo de control estará conformado por individuos no seleccionados de forma aleatoria y que posean características semejantes a los individuos integrantes del grupo experimental. Los grupos se comparan luego de aplicado el estímulo al grupo experimental para tomar mediciones de los resultados y analizar si el estímulo tuvo algún efecto sobre la variable dependiente. Un diagrama del diseño del experimento se puede ver a continuación, donde puede observarse que no existe una asignación aleatoria ni emparejamiento entre los grupos de control y experimental.



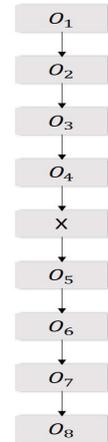
La principal fuente de error proviene de la no aleatoriedad de los grupos de control y experimental. Dependiendo del grado de no correspondencia, es decir del nivel en que un grupo se diferencie del otro, las características más diferenciadas podrían introducir efectos que tenderían a confundirse con los producidos por la variable independiente, desvirtuando las mediciones sobre la variable dependiente, quien es el objeto de la investigación.

A pesar de los inconvenientes que podrían presentarse en este diseño, los problemas de desviación producto de características muy diferentes entre los individuos integrantes de los grupos podrían resolverse aplicando el análisis de covarianza sobre los resultados finales obtenidos, siempre y cuando la variable de covarianza se escoja de manera adecuada.

#### B)Diseño cuasiexperimental de series cronológicas

En algunas investigaciones es necesario comprender el efecto a mediano y largo plazo del estímulo experimental cuando es aplicado sobre una población o grupo que no es posible escoger o seleccionar de forma aleatoria. En esta situación se realizan varias mediciones de la variable dependiente y se intercala el estímulo experimental entre ellas en el grupo experimental. Nótese que no se cuenta con un grupo de control de comparación, sino con un solo grupo al cual se le realizan varias mediciones antes y después de aplicado el estímulo experimental. A esto se le denomina diseño de series cronológicas.

Los posibles resultados entre dos medidas se pueden observar en la figura 5. Según lo que indica la bibliografía clásica, los resultados de introducir un estímulo en X es una serie cronológica de mediciones que abarca desde  $O_1$ -1- $O_8$ . En el punto donde se aplica el estímulo en la serie puede observarse un comportamiento consistente, salvo en el caso D, donde la diferencia en el efecto es apreciable, aunque los comportamientos indican que el efecto tiene variaciones importantes, donde A y B muestran un máximo en la transición y en H, F y G el comportamiento es aleatorio por lo que no se puede justificar qué lo produce.



La característica con mayor dificultad de establecer es el factor historia, que representa la influencia de condiciones externas al experimento ocurridos durante el desarrollo del mismo. Otro factor es el método de análisis escogido para conocer el efecto de la variable independiente sobre los individuos integrantes del grupo experimental. La práctica indica que el análisis de varianza, el cual promedia los valores antes y después de aplicado el estímulo, es el más adecuado para el análisis de resultados.

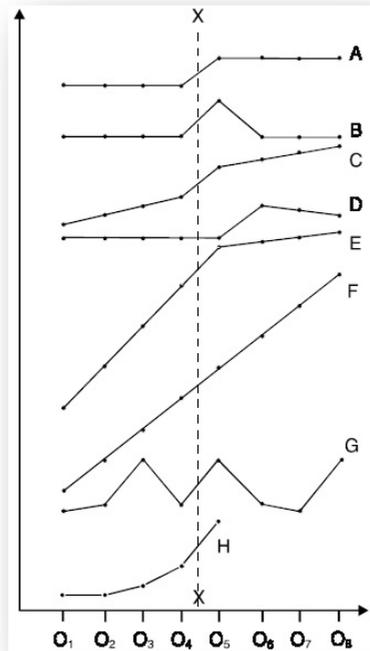


Figura 5. Serie cronológica de cuatro observaciones antes y después del estímulo experimental

## Diseños no experimentales

A diferencia de los diseños experimentales y cuasi experimentales, en los diseños no experimentales el investigador no ejerce ningún control sobre la variable independiente y tampoco estructura grupos de estudio. En estas investigaciones la variable independiente ya ha producido un efecto cuando se comienza el estudio, lo que indica que ya ha sucedido algo, que se ha producido alguna consecuencia sobre el grupo de estudio; partiendo de este punto el fenómeno ocurrido será identificado como la variable independiente, como por ejemplo el consumo de gaseosas, para la elaboración de un proyecto de estudio entorno a él. En dicho proyecto el investigador pretende describir tanto la variable independiente como sus efectos sobre otras variables que vendrían a ser dependientes de su estímulo, como, por ejemplo, propensión a la presencia de diabetes y problemas circulatorios.

Debido a la configuración de estos estudios, donde la incidencia de la variable independiente ya ha ocurrido, los diseños no experimentales también reciben el nombre de investigaciones ex post facto (después del hecho).

Algunos ejemplos podrían ser

- Estudio de las conexiones cerebrales de los taxistas expertos de la ciudad de Lima.
- Efectos de las redes sociales sobre los adultos jóvenes y la capacidad de ejecutar tareas de alta concentración.

Veamos las características de las investigaciones con diseño no experimental

- El investigador escoge los efectos de un estímulo que puede ser identificado y que es posible observar para poder hacer una regresión en tiempo e identificar relaciones y las posibles causas.
- Es adecuado cuando no es posible realizar experimentaciones sobre el grupo objeto de estudio por diversas razones, que pueden ir de razones económicas a razones éticas.
- Los estudios ex post facto brindan información útil sobre cómo se asocian las causas con las consecuencias de un problema, en qué orden y en cuales condiciones; para ello se apoya en técnicas estadísticas como la correlación parcial y la regresión múltiple.
- La debilidad de este diseño es que no se puede filtrar con certeza los factores causales de un fenómeno.

### Tipos de experimentos de diseños no experimentales

Entre los diseños no experimentales se encuentran la encuesta social, el estudio de cohortes, el estudio de casos, la observación estructurada, la investigación evaluativa, la investigación-acción participativa. La encuesta social es una de las herramientas de los diseños no experimentales más utilizadas y será tratada en un capítulo aparte. A continuación, algunos tipos de encuestas.

## Encuestas según la temporalidad: seccionales y longitudinales

### Encuestas seccionales

Las encuestas que se realizan en un período corto de tiempo y que intentan reflejar un fenómeno en un instante determinado, como la tendencia de un electorado a votar por un candidato, se le denominan encuestas seccionales o síncronas y un caso de ellas es la encuesta ómnibus, que son breves cuestionarios aplicados al mismo tiempo sobre una población específica.

### Encuestas longitudinales

Son aquellas encuestas que se emplean en estudios que se realizan a lo largo de un tiempo prolongado. Las encuestas longitudinales o diacrónicas, pueden basarse en datos pasados, para el cual se tendría entonces un estudio retrospectivo. Un ejemplo sería el estudio del historial médico de una cierta población.

Cuando el estudio se basa en varias encuestas aplicadas durante un período prolongado de tiempo se está en presencia de estudios prospectivos. Entre ellas se encuentran las encuestas del método Delphi y las encuestas de los estudios de panel.

### Desarrollo del proceso al momento de aplicar una encuesta social

El proceso para aplicar una encuesta social se extiende desde la concepción del proyecto hasta su ejecución y posterior socialización. Dicho proceso cumple con las siguientes etapas.

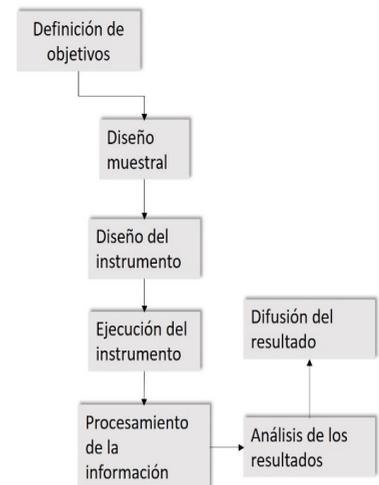
a) Se enuncia un problema de investigación y se circunscribe a su entorno inmediato de acción, definiendo los antecedentes o marco teórico.

b) Se formulan los objetivos o las hipótesis que definen el problema. La formulación de los objetivos debe ser descriptiva y enfocada a un logro. La hipótesis debe ser descrita como la relación entre variables independientes y dependientes o como una hipótesis estadística.

c) Se determina la envergadura de la encuesta. Lo que comprende:

c.1) Definición de la población objetivo. Marco geográfico y poblacional.

c.2) Definición de la población en términos demográficos (sexo, edad, nivel educativo, etc.). Se debe considerar la composición poblacional dependiendo si la encuesta es descriptiva o explicativa.



c.3) tiempo que tomará la ejecución de la encuesta, fecha y período.

Dependiendo del criterio de la investigación, si la envergadura de la encuesta es total o parcial. Si es total se debe recoger información en toda la población o intervalo de estudio. Si es parcial se debe definir el tamaño y el tipo de muestra a tratar, como los posibles inconvenientes de logística (personas que no aceptan responder la encuesta, direcciones erróneas, números de contacto falsos, etc.).

d) Determinar las variables a utilizar. Dichas variables deben ser medibles, simples y suficientes para cubrir los objetivos del estudio. Además, deben ser capaces de permitir la construcción de variables complejas.

e) Construir los instrumentos que serán utilizados para la obtención de la información, por ejemplo, cuestionarios con preguntas abiertas o cerradas. Si se requiere medir una variable que requiera cierta gradación en su aceptación o rechazo, la encuesta debe reflejar la técnica más acorde a ello.

f) Se realizan pruebas de los instrumentos en grupos reducidos que posean las características que se pretenden medir en la población de estudio.

g) Se traslada el diseño a pruebas reales. Los entrevistadores deben estar capacitados para la labor asignada in situ. Se debe realizar controles periódicos sobre la recolección de datos ejecutada por los entrevistadores, a fin de asegurar que se realiza la recolección de manera adecuada. Por lo general se supervisan cierta cantidad de encuestas de forma aleatoria.

h) Se realiza control sobre la calidad de la información recolectada, revisando la coherencia de las respuestas, la consistencia de las mismas y su completitud. De allí se decide cuáles entrevistas eliminar, reprogramar o en dado caso reformular. Aquí ya está definido el sistema de interpretación de resultados, con sus codificaciones respectivas.

i) Se procede al análisis de resultados y su posterior interpretación, siguiendo un esquema de análisis ya preparado para ello. Por lo general comprenden tratamientos estadísticos de los resultados, quienes general las respuestas cuantitativas tanto a las preguntas como los objetivos de la investigación mediante gráficos, tablas de frecuencia, etc.

En las investigaciones cuantitativas se busca comparar resultados, encontrar patrones de incidencia, relacionarlas con la problemática estudiada y pronosticar probables comportamientos futuros, explicar la relación entre variables que definen una hipótesis, etc.

El procesamiento de datos es fundamental para comprender los resultados del estudio y ellos serán indispensables para explicar y exponer la culminación de la investigación.

El análisis de los resultados permite comprender la importancia del estudio

j)Presentación del informe final, el cual consta la exposición del problema estudiado, sus antecedentes y marco teórico, la muestra o población objeto de estudio, el período de tiempo en el cual se ejecutó la labor de campo, el camino metodológico seguido, las dificultades que se presentaron y los resultados que arrojaron los análisis hechos sobre el objetivo, las conclusiones a las que se llegaron y la documentación de los instrumentos utilizados en la recolección de datos.

# CAPÍTULO 4

## Recolección de datos sociales

La recolección de datos es una de las tareas que requiere tanto especial rigor como esfuerzo durante el desarrollo de un proyecto de investigación. Esta inicia con el empleo de estrategias y herramientas o métodos según las necesidades del diseño de investigación. El diseño de investigación es el que condiciona la metodología a utilizar al momento de realizar la recolección de datos; dependiendo del tiempo disponible, de la financiación y el recurso humano, los métodos que pueden utilizarse son la observación y las entrevistas, los formularios y las encuestas.

### **La encuesta Social**

La encuesta es uno de los recursos más empleados en la investigación social cuantitativa, esto es motivado a su flexibilidad para recolectar información de una población o de una muestra de ella. Las encuestas hacen uso de preguntas, bien sean orales o escritas para que los individuos seleccionados según los requerimientos del proyecto de investigación, puedan responder a las interrogantes planteadas y que busquen describir o explicar el fenómeno estudiado.

Como se indicó, las encuestas poseen un grado destacable de flexibilidad, por lo que permiten recoger información de muy diversas índoles y por ello sea de mucha utilidad en múltiples áreas del conocimiento como en marketing, ciencias políticas, psicología de masas, sociología, educación, demografía entre otros. Según sea el caso de estudio, la información que se requiera extraer de la población o muestra puede estar clasificada de la siguiente manera:

- Conductuales: tendencias políticas, actividades recreativas, actividades sociales, ética tecnológica, etc.
- Actitudinales: preferencias afectivas, liderazgo, causas político-sociales, posición religiosa, etc.
- Demográficas: edad, sexo, altura, peso, domicilio, genero, estado civil, etc.
- Socio-económicas: escolaridad, movilidad social, ingresos, salarios, ocupación.

Estos tipos de información se pueden obtener utilizando cuestionarios con un procedimiento determinado, el cual responde a los objetivos de la investigación. Estos procedimientos pueden ser mediante:

- Entrevistas personales con los individuos seleccionados para el estudio.
- Entrevistas telefónicas Online.
- Entrega de cuestionarios para ser rellenados sin la presencia de un entrevistador y devueltos en un período de tiempo determinado.

Como es natural pensar, cada uno de estos procedimientos tiene sus ventajas y desventajas. Las entrevistas telefónicas permiten un desarrollo ágil al momento de completar los formularios, pero tienen la desventaja de requerir que los entrevistados tengan un teléfono a disposición y es más sencillo para las personas negarse a realizar la encuesta. La realización de una encuesta en presencia de un entrevistador podría cohibir a las personas a ser completamente francas con sus respuestas, sin embargo, el poder aclarar dudas sobre los aspectos de una pregunta permite una precisión mayor en las respuestas. Los formularios entregados para ser completados por las personas en su hogar tienen la desventaja de que algunas personas no sean tan rigurosas en las respuestas, pudiendo quedar incompletas o ambiguas; la ventaja es que las personas, al no sentirse presionado o cohibido por un entrevistador podrían pensar mejor sus respuestas y ser más objetivas en ellas.

En lo que respecta a la orientación de los objetivos de la investigación social cuantitativas, las encuestas pueden ser de dos tipos, las encuestas descriptivas y las encuestas explicativas.

### **La encuesta descriptiva**

Criterio Porcentaje (100 %)

Doctrina política y económica donde los trabajadores controlan los medios de producción.	45
Es la fase anterior al comunismo.	15
Es sistema político de partido único.	22
Es un estado de carácter totalitario.	18

Este tipo de encuesta busca como objetivo describir cómo una o más variables se distribuyen sobre una población o muestra que es objeto de estudio, calculando valores como la media central y la dispersión que se observe sobre dicha población o muestra. Además, es posible realizar los mismos cálculos sobre un subconjunto de ellos; esto permite a su vez comparar la forma en que se distribuyen los valores de las variables tanto en todo el conjunto como en algún segmento de él. Un ejemplo de encuesta descriptiva podría ser sondear el conocimiento de un conjunto de X personas acerca del socialismo.

- Según el criterio del conjunto de personas X, valorado en porcentaje (%), el socialismo consiste en lo descrito en la tabla 2.

**Tabla 2. Porcentaje según cada criterio**

<b>Criterio</b>	<b>Porcentaje (100 %)</b>
Doctrina política y económica donde los trabajadores controlan los medios de producción.	45
Es la fase anterior al comunismo.	15
Es sistema político de partido único.	22
Es un estado de carácter totalitario.	18

X, valorado en porcentaje (%), el socialismo consiste en:

**Tabla 3. Criterios según el género**

- Nivel de escolaridad del conjunto de personas X, valorado en porcentaje (%).

<b>Criterio</b>	<b>Porcentaje hombres (100%)</b>	<b>Porcentaje mujeres (100 %)</b>
Doctrina política y económica donde los trabajadores controlan los medios de producción.	22	35
Es la fase anterior al comunismo.	28	32
Es un sistema político de partido único.	20	17
Es un estado de carácter totalitario.	30	16

**Tabla 4. Escolaridad**

<b>Escolaridad en años</b>		
<b>Hombres</b>	<b>Mujeres</b>	<b>Promedio</b>
7.2	8.3	7.8

Como muestra el ejemplo, crear subconjuntos con variables de tipo factual (escolaridad, sexo, genero, edad, salario, etc.), como la escolaridad en los subconjuntos de hombres y mujeres del ejemplo, le permite al investigador comparar cómo se comportan las variables de manera general en todo el grupo y de manera parcial entre estos subconjuntos, estableciendo las variaciones que se detectan para poder luego formular conclusiones. Es necesario en las encuestas descriptivas elaborar formularios que contemplen estas variables que ayudan a definir subconjuntos, pero debe indicarse también que algunas variables no factuales pueden incorporarse para definir subconjuntos. Así como la escolaridad define al subconjunto de hombres y mujeres, lo mismo puede hacer la variable ideología en el mismo sentido, si es operativizada adecuadamente, los diferentes niveles de dicha variable podrían funcionar como categorías utilizadas como medidas de comparación.

Como puede notarse, el objetivo de las encuestas descriptivas es comparar resultados, y para poder lograrlo, el conjunto sujeto de estudio debe ser lo suficientemente diverso para que permita la categorización de subconjuntos y que las respuestas de esos subconjuntos a los objetivos de investigación permitan establecer diferencias observables. Pero cabe destacar que la diversidad no es una característica introducida al conjunto sujeto de estudio, sino que debido a esas características propias es que se utiliza a ese conjunto como sujeto de estudio. Por ello el investigador debe contextualizar en el momento de la elaboración de los cuestionarios, aquellos factores que puedan servir como comparadores. Además, para lograr una comparación con significado, deben existir normas que permitan la evaluación de los resultados de algunas comparaciones. Por ejemplo, si se va a medir la satisfacción de las personas con respecto al servicio del agua potable, es recomendable que se compare y referencie con otros servicios públicos como la telefonía o el servicio de electricidad.

### **Proceso de análisis de los datos de las encuestas descriptivas**

Cuando ya se ha completado el proceso de recolección de datos se procede a su adecuación para la etapa siguiente que es el análisis de datos. Para ello se debe cumplir la siguiente ruta.

a) Se tabulan el total de los datos de cada aspecto del problema consultado según los requerimientos presentes en los objetivos de investigación, sacando una descripción cuantitativa o una medida estadística sobre todos los datos.

b) De igual manera se tabulan los datos obtenidos de los subconjuntos consultados y que fueron seleccionados debido a la relevancia para los objetivos de la investigación de las descripciones que podían resultar de la recolección de datos de dichos segmentos.

c) Según su importancia, sobre el conjunto y subconjuntos de individuos objetos de estudio, se aplican procesos estadísticos para la depuración de resultados (desviación estándar, media aritmética, etc.). En este punto también se puede realizar operaciones sobre los resultados como la elaboración de índices, perfiles, etc.

d) Los datos son comparados con otros obtenidos en investigaciones similares para una mejor evaluación de resultados.

e) Se complementan los resultados con materiales cuya procedencia no tenga un carácter cuantitativo, como aclaratorias, anexos, comentarios, etc. Que permitan explicar con mayor claridad lo que de otra manera solo serían informes de contenido técnico y que no ayudan a la comprensión de los logros



obtenidos y el significado que tienen.

## **La encuesta explicativa**

Este tipo de encuestas busca dar explicación, en general mediante una hipótesis, de cómo se produce un fenómeno, El fenómeno estaría representado por una variable dependiente, en función de su relación con uno o más estímulos, que representarían a las variables independientes o factores causantes.

Dependiendo del problema a abordar mediante el análisis de esta herramienta, se podrían tener al menos dos tipos de análisis: 1) el análisis mediante la especificación y la explicación de la relación entre variables, y 2) la interpretación del fenómeno estudiado.

Como se pudo ampliar en el capítulo anterior, la capacidad de identificar un factor o variable independiente como causante de las variaciones de otra, denominada variable dependiente, está ligada al control que se ejerce sobre otros factores o variables que pudiesen añadir distorsión en las mediciones. La encuesta explicativa trata de emular el control e identificación que se ejerce sobre las variables externas en un diseño experimental, pero en un escenario natural. Para ello, en vez de aislar y manipular una variable independiente para monitorear su impacto sobre la variable dependiente, el investigador busca encontrar en el ambiente en el que pretende desarrollar su investigación, las consecuencias o efectos de la que será su variable independiente. Ello representa un problema a resolver por la muy probable presencia de muchas variables que pudiesen afectar de una u otra manera el fenómeno que se pretende estudiar. Por tanto, el investigador requiere reducir su universo de interés y focalizarse en la que será la muestra representativa a la que aplicará la encuesta explicativa donde se podrán excluir las otras fuentes de distorsión en la medición del fenómeno a estudiar. En otras palabras, en la encuesta explicativa, se confía en que se logra identificar a la variable independiente restringiendo la heterogeneidad de la población de la cual se seleccionará la muestra a la cual se aplicará el estudio.

Para ilustrar este procedimiento, se puede recurrir a un ejemplo donde se pretende confirmar la hipotética relación entre el éxito del populismo y el descontento de segmentos poblacionales que no han cursado una educación universitaria. El hecho de centrar el estudio en un universo donde los que han cursado una carrera universitaria son por estudios más preparados, como mayores habilidades para ejercer profesiones y lucrarse de sus conocimientos e intelectualmente más educados pueden identificar todos los peligros que entrañan las tácticas de políticos populistas, elementos que una persona con educación básica, posibilidades económicas más limitadas y un intelecto menos entrenado podría creer con mayor facilidad en las promesas de reivindicación social populista. Teniendo en cuenta esta línea de pensamiento se descartan todos aquellos sectores que no cumplen el perfil señalado, restringiendo a unos pocos segmentos demográficos, a ciertas edades como las que pueden votar, etc.

Motivado a las condiciones restrictivas del universo elegible, es usual que la muestra deba ser grande para contar con la suficiente cantidad de casos necesarios para organizar categorías, en el ejemplo serían categorías socioeconómicas, para poder aplicar análisis estadísticos (en especial cuando se requiere para realizar pruebas de significado) sobre ellas y poder explicar cómo ocurre el fenómeno estudiado en la realidad objetiva.

## La encuesta social

Existen puntos de encuentro entre el diseño de una investigación social y el diseño de una encuesta social motivado a que ambas cumplen procedimientos similares en el desarrollo de sus etapas. En la investigación social cuantitativa se sigue el siguiente esquema:

- a) Un primer acercamiento al problema que se quiere abordar el cual parte de una idea de investigación.
- b) La preparación de un plan o proyecto de investigación donde se indica el proceso a seguir para abordar el problema. Estas tareas son las siguientes:
  - b.1) Ideas y planteamiento de un problema de investigación.
  - b.2) Revisión del soporte teórico, bases y fundamentos: Antecedentes del problema.
  - b.3) Alcance del proyecto: Objetivos de la investigación.
  - b.4) Diseño de la metodología.
  - b.5) Diagrama de trabajo: recolección, procesamiento y análisis de datos.
  - b.6) Elaboración del reporte final.
- c) Se desarrolla el proceso de ejecución del plan. Esto es la instrumentalización del plan o proyecto. Con las siguientes etapas:
  - c.1) Selección de la población objetivo, y si es pertinente se toma una muestra de la misma.
  - c.2) Se acota el tamaño de la muestra y las características de la misma.
  - c.3) Selección de los instrumentos de recolección de datos (encuesta, formularios, cuestionarios, etc.).
  - c.4) Ensayos de efectividad de los instrumentos; en caso de introducir modificaciones se realizará el rediseño de los instrumentos a utilizar.
  - c.5) Entrenamiento de los equipos de apoyo y de recolección de datos mediante pruebas de campo.
  - c.6) Selección y depuración de datos (descartando datos incompletos, ambiguos, falsos, etc.).
  - c.7) Estructuración de los datos depurados en un formato ya prediseñado para su manipulación.
  - c.8) Acondicionamiento del modelo de análisis a utilizar según lo definido en los objetivos del proyecto (tablas estadísticas, histogramas, cálculos estadístico-descriptivo, regresiones, correlaciones, etc.).



- c.9) Elección de paquetes de procesamiento estadístico (Excel, SPSS. etc.). Ponderar el uso de procesos estadísticos de manera manual si la población y los datos no son muy grandes.
- c.10) Organización de la información procesada para la elaboración del informe.
- c.11) Elaboración de las partes del informe: introducción, antecedentes, capítulos, anexos, soportes.
- c.12) Preparación de la bibliografía pertinente al proyecto según las normas de citación establecidas previamente (APA, IEEE, Vancouver, etc.).
- c.13) Lectura interpretativa de los análisis hechos a la información obtenida.
- c.14) Primer borrador del informe.
- c.15) Informe final.

Dada la naturaleza de la encuesta social se tomará aquellas tareas más representativas de forma resumida, añadiendo algunas que no están en la lista.

### **Determinación de la muestra**

Cuando el investigador se encuentra con que el conjunto de individuos que podrían ser objeto de estudio es muy amplio, debe mediante técnicas probabilísticas extraer un grupo representativo de dicho conjunto. Para poder exponer mejor el procedimiento necesario para lograr escoger este grupo representativo se requiere manejar ciertos términos básicos como lo son la población o universo, el marco de muestreo, la muestra, la unidad de muestreo, la fracción de muestreo y la representatividad de la muestra.

### **Población**

Representa al total de individuos o unidades que integran el conjunto de personas o elementos que serán objeto de estudio y dónde se abordará el problema de investigación presente en el proyecto de investigación. Si por ejemplo se estudiará el populismo, serán personas con capacidad de votar, hombres y mujeres mayores de 18 años, etc. La delimitación de la población se ejecuta según los requerimientos del proyecto de investigación; por ejemplo, en el caso del estudio del populismo, se restringirá el estudio a personas con poder de voto. La bibliografía representa a la población con la letra N.

### **Marco de muestreo**

Se define como una lista de todos los elementos que constituyen a la población objeto de estudio y de quién se extrae la muestra. Los elementos del marco de muestreo pueden ser individuos, pero también pueden ser instituciones, empresas, hogares o cualquier otra unidad que pueda ser objeto de investigación.

### **Muestra**

Es la que identifica al conjunto de individuos o elementos que serán objeto de estudio en el desarrollo

del proyecto investigación y que fueron seleccionados mediante algún procedimiento probabilístico. Para representar a la muestra se utiliza la letra  $n$ .

### **Unidad de muestreo**

Se refiere al elemento mínimo que representa a la muestra, pudiese ser individuos como unidad más simple, o unidades más complejas como instituciones, universidades, provincias, etc. La unidad debe estar identificada en el marco de muestreo.

### **Fracción de muestreo**

Es el cociente entre la muestra y la población. Esta nos indica el grado proporcional de la muestra con respecto al universo de elementos del que deriva. Se simboliza con la letra  $f$  y la relación se escribe como  $f = n/N$ . Por ejemplo, si la muestra está conformada por 400 individuos de un universo de 12000 personas la relación se escribe como  $f = 400/12000$ , de donde  $f = 1/30$ . Es decir, que la muestra es la trigésima parte del universo posible.

### **Representatividad de la muestra**

Es el grado en el que la muestra reproduce aquellas singularidades que caracterizan a la población de la que proviene. A mayor representatividad mayor es la facilidad con la que cualquier unidad presente en la población puede ser parte de la muestra.

### **Tipos de muestras**

Existen dos tipos fundamentales de muestras, las probabilísticas y las no probabilísticas. Las muestras probabilísticas son escogidas de forma aleatoria, donde cada unidad de la población tiene una probabilidad distinta de formar parte de la muestra.

Las muestras no probabilísticas no son escogidas de forma aleatoria, por tanto, existe direccionalidad en su selección.

### **Tipos de muestras probabilísticas**

Dentro de las muestras probabilísticas existen los siguientes tipos clasificados por su característica principal.

#### **Muestra aleatoria simple**

En este tipo de muestra, todas las unidades de la población poseen la misma probabilidad de ser parte de la muestra. En 500 unidades ya tabuladas, la elección de los elementos que pertenecerán a la muestra se escoge al azar, mediante una tabla de valores al azar o mediante alguna función informática de tipo random ejecutada desde un computador. El único condicionante es el tamaño que tendrá la muestra.

## **Muestra estratificada proporcional**

Cabe señalar que, dentro de esta categoría existes varios subtipos de muestras estratificadas, pero el proceso general indica que, antes de seleccionar la muestra, las unidades que componen a la población se distribuyen en estratos, de donde se selecciona las unidades de la muestra, manteniendo la proporcionalidad escogida con la fracción de muestreo  $f$ . Por ejemplo, si en un universo de 500 jugadores de futbol, 400 son diestros y 100 siniestros y la muestra a escoger es de 25 jugadores, entonces  $f = 25/500$ , donde  $f = 1/20$ . Para mantener la proporcionalidad deben escogerse 20 jugadores diestros y 5 siniestros.

Ya identificada la cantidad de unidades que deben seleccionarse de cada estrato, la selección de dichas unidades se hace de forma aleatoria como se haría en la selección de la muestra simple o mediante un intervalo de selección.

## **Muestra sistemática**

Se obtiene dividiendo en un número aleatorio de secciones o intervalos a la población, se escogerá un número al azar que esté comprendido entre el 1 y el tamaño de cada intervalo para luego a partir de ese elemento seleccionar cada integrante de la muestra. Por ejemplo, si se tiene una población de 1000 unidades y se requiere una muestra de 100 unidades, se procede a dividir a la población en 100 partes aleatorias de 10 unidades. Como el rango de cada intervalo es de 10 unidades, se escogerá un número aleatorio del 1 al 10; si el número resultante del azar es 4, se comienza con la unidad 4 de cada intervalo hasta completar la muestra establecida, si son 100 unidades se escogerá al cuarto individuo de cada intervalo, es decir las unidades 4, 14, 24, y así sucesivamente hasta completar la muestra.

Cuando la muestra no tiene simetría con el tamaño de la población, lo que quiere decir que no es múltiplo de ella, el cociente  $N/n$  se aproxima al valor más próximo por encima o por debajo. Por ejemplo, si la población es de 235 unidades, de la cuales 20 serán tomadas de muestra, entonces  $235/20 = 11.75$ . Se redondea a 12 y se realiza el mismo procedimiento visto en el caso anterior.

## **Muestra de conglomerados**

Se utiliza cuando se requiere hacer estudios sobre poblaciones que se encuentran en una zona geográfica extensa o que se presentan mediante agrupaciones naturales que identifican de forma adecuada a la población. Si existen estas características, se puede seleccionar algunas de estas agrupaciones como objeto de estudio.

## **La selección de la muestra toma el siguiente procedimiento**

-Se identifica la característica que permitirá agrupar a la población en conjuntos de forma intensiva -es decir, que no quede ningún individuo sin pertenecer a un conjunto- y que dichos conjuntos no se solapen entre sí -condición disjuntiva-.

-Se escogen de forma aleatoria algunos conjuntos formados.

-Se toma al azar las muestras de cada conjunto seleccionado, siguiendo un camino aleatorio simple o mediante el sistema de muestra sistemática.

Uno de los criterios como se mencionó al principio es el geográfico, por ejemplo, si el investigador llega a la conclusión de que los enfermos de diabetes son proporcionales en cada estado o provincia de un país, puede seleccionar algunos de esas provincias o estados para realizar la investigación. Otro criterio es el de grupo naturales, como los que se observan en los paralelos o cohortes de derecho romano de varias universidades. En ambos casos se puede aplicar los criterios para seleccionar una muestra por conglomerados.

En el caso particular de las cohortes de matemáticas, si se tienen 25 secciones de 20 alumnos de derecho romano y se quiere manejar una muestra de 50 individuos, se tiene que  $f = 50/500 = 1/10$ , por lo que se pueden seleccionar  $2 (2 \times 25 = 50)$  cursos de forma aleatoria para formar la muestra.

En la realidad los conjunto o conglomerados no se presentan con la uniformidad con la que se ha tratado los ejemplos para emplear una muestra basada en conglomerados. En la realidad los conjuntos no son homogéneos y por ello requiere un procedimiento más elaborado como sería la selección de la muestra mediante la probabilística proporcional al tamaño de los conjuntos identificados. Sin embargo, en general la muestra por conglomerados permite la economía de recursos, como costes en desplazamiento, producción de material entre otros.

### **Tamaño de la muestra**

En la investigación social es importante conocer cual es el tamaño de la muestra que cumple con la mayor eficiencia en función a la calidad de los resultados buscados. Conocer el tamaño ideal de la muestra depende de la disposición de información con el suficiente significado para poder realizar cálculos en base a ellos y de esta manera conocer el valor del tamaño adecuado para la muestra. Entre la información necesaria se encuentran los siguientes:

- 1) Valor estadístico principal a medir que será el eje del análisis derivado de la encuesta aplicada (mediciones aritméticas, proporcionalidades, etc.)
- 2) Tamaño de la población. Si es posible conocer todos sus integrantes -finita- o no -infinita-.
- 3) Profundidad del análisis (análisis sobre la totalidad de los resultados, análisis sobre segmentos o subgrupos).
- 4) Error aceptable en los valores obtenidos de la muestra.
- 5) Probabilidad que el error muestral observado tenga la magnitud o valor estimado.
- 6) Recursos de los que se dispone (financieros, equipos, humanos).

Cuando se dispone de la información necesaria como la descrita arriba, se pueden aplicar fórmulas que nos ayuden a conocer el tamaño de la muestra que es más adecuado para un proyecto de investigación. Cuando se conocen los datos requeridos como valores promedio en poblaciones donde no se contabilizan todos sus miembros, se puede aplicar la fórmula.

$$n = \frac{z^2 \cdot p \cdot q}{e^2} \quad (1)$$

Donde  $z$  es el nivel de confianza que puede conocerse mediante tablas estadísticas,  $p$  es el valor de la probabilidad de éxito o proporcionalidad esperada. Esto depende de la información que se tenga de investigaciones previas, si esa información no se conoce se estima darle un 50% de éxito, y  $q$  es la probabilidad de fracaso, y al igual que  $p$ , si no se tiene alguna información previa que ayude a conocer dicha estimación, se le coloca una probabilidad de 50%. Por último,  $e$  es la precisión o error máximo permitido y lo fija el investigador.

Por ejemplo, si en un estudio requiere conocer la cantidad de autos con combustión Diesel que puedan existir en una ciudad con un error no mayor al 4% y con una confiabilidad del 96%. Entonces al no conocer el tamaño de la población tendremos que sustituyen en la fórmula anterior no arroja los siguiente.

$$n = \frac{(2,05)^2 \cdot 0,5 \cdot 0,5}{(0,04)^2} \quad (2)$$

$$n \approx 656$$

El ejemplo nos indica que se necesitarían 656 individuos seleccionados al azar para llevar a cabo la investigación.

### **Muestras no probabilísticas**

Como se mencionó en un párrafo anterior, las muestras no probabilísticas no se escogen al azar, pues su finalidad le permite tener direccionalidad que responde a la búsqueda de características específicas seleccionadas por el investigador al momento de escoger la muestra. Entonces, la muestra escogida responde a un sesgo impuesto por el investigador según su criterio personal y por tanto no se pueden aplicar técnicas estadísticas inferenciales ni es posible calcular el error en los valores. Existen tipos de muestras no probabilísticas, entre ellas están

- Muestra de sujetos voluntarios: se utilizan sobre todo en medicina, donde sujetos se ofrecen de voluntarios para probar y monitorear los efectos algún medicamento. con este tipo de muestra no se puede inferir si los individuos son representativos de alguna población.

- Muestras por conveniencia: son aquellas que se selección según la finalidad de investigación sobre un grupo de estudio, por ejemplo, pacientes con cáncer. En estos casos el hospital no atiende a toda la población.

- Muestras por cuotas: por lo general el investigador escoge de forma proporcional los participantes de la muestra. En un universo de 100 personas elegidas entre 18 y 19 años, 50 serán de 18 años y los otros 50 serán de 19 años. Dentro de los escogidos con 18 años, habrá 25 hombres y 25 mujeres, igual en el grupo de 19 años. Es decir, se guarda una proporción simétrica en los grupos seleccionados.

## El cuestionario

El componente central de una encuesta es el cuestionario. Este no se construye según algún manual preexistente, sino que se elabora de acuerdo a la experiencia del investigador quien reconocer los elementos que más se adecuan a la naturaleza del proyecto de investigación que desea desarrollar. Sin embargo, se pueden reconocer ciertas pautas generales en su elaboración que pueden servir de referencia. De estas tratamos a continuación.

a) Cuando se identifica el problema de investigación y se fijan los objetivos del estudio, las preguntas del cuestionario deben responder al planteamiento de ellas.

b) Aunque no existe una metodología para conocer la cantidad de preguntas que debe contener el cuestionario, es necesario plantear en un primer acercamiento la cantidad de preguntas que se consideren necesarias para abarcar los objetivos de la investigación. Luego en una segunda revisión se realizará la depuración de aquellas preguntas que, aunque parecían importantes no evalúan adecuadamente algún aspecto de los objetivos de investigación.

c) Se recomienda que el cuestionario inicie con preguntas de índole general, que no revistan mayor complejidad y que puedan estar al alcance de cualquier encuestado, esto fomenta un ambiente favorable donde el entrevistado se sienta cómodo con el entrevistador.

d) Luego de iniciar con las preguntas generales se prosigue en profundidad con las preguntas específicas del cuestionario que abordan temas asociados a la investigación. Por ejemplo, el cuestionario podría contemplar una pregunta general como ¿Qué opinión tiene sobre el alcalde de su provincia? Entonces después de esta pregunta general podría formular una pregunta más específica como ¿Qué opina del desempeño del alcalde con respecto a la ejecución de las políticas de aseo urbano? Esto le brinda un contexto al entrevistado para emitir sus respuestas.

e) Se recomienda que las preguntas guarden una secuencia lógica que emule a una conversación natural, avisando al encuestado cuándo se tratará un tema distinto al que habían abordado hasta ese momento. Por ejemplo, al momento de abordar otro tema decirle al encuestado “ahora hablaremos del siguiente tema” o “queremos pasar a otro tema”

f) No formule preguntas que puedan direccionar la respuesta del encuestado. Por ejemplo, no preguntar “¿siente que es muy incómodo llevar mascarilla?” lo que podría llevar al encuestado a responder afirmativamente de manera inmediata.

g) Evite las preguntas que pudiesen ser incómodas para el encuestado y que pudiese inducir una respuesta no sincera. La pregunta “¿aprueba usted el aborto?” podría reformularse como “existen personas que están a favor y otros en contra del aborto; ¿Qué opina al respecto?”

h) Evite formular preguntas vagas que no puedan ser medibles en el análisis posterior, como aquellas que contienen expresiones como “ha menudo usted...” “por lo general...” “con cual frecuencia...”, etc.

i) El cuestionario debe estar estructurado en tres secciones:

1) Verificación del entrevistado: permite conocer si el encuestado pertenece al grupo de interés de la investigación. Por ejemplo “¿Posee carnet de conducir?” “¿está inscrito en el registro de votantes?”.

2) Clasificación factual y demográfica: preguntas que estratifican al encuestado en un segmento de la población (edad, sexo, escolaridad). Estos datos suelen provocar inquietud en el encuestado por lo que se recomienda ubicarlas en la sección final del cuestionario.

3) Preguntas específicas: son por diseño la sección más extensa del cuestionario, pues en ella se abordan las preguntas que dan respuesta a las interrogantes de investigación. Pueden ser abiertas o cerradas.

j) Dependiendo de los objetivos de la investigación y del análisis que será aplicado el cuestionario utilizará alguno de estos tipos de preguntas o una combinación de ellas, que se muestran a continuación.

1) Preguntas cerradas. Le permite al encuestado escoger una respuesta entre varias alternativas para de esta manera escoja aquella que mejor se aproxime a su criterio de pensamiento. Entre ella tenemos

-Dicotómicas: “¿Le han hecho algún tipo de cirugía?”

Sí\_ No\_

-Selección Múltiple: “¿usted cómo definiría su religiosidad”

Católico practicante\_ Católico no practicante\_

Agnóstico\_ Budista\_

Protestante practicante\_ Protestante no practicante\_

Ateo\_ Creyente\_

Poseen las siguientes ventajas

- Los resultados tienden a tener una interpretación más uniforme.
- Son fáciles de administrar, procesar y analizar.

Tienen como desventaja que tienden a ser rígidas, con un nivel de profundidad superficial y por ende adolecen de precisión al reflejar la real opinión de los encuestados.

2) Preguntas abiertas. El entrevistado puede contestar indicando una o más respuestas o el orden de importancia de un ordenamiento de opciones.

- Pregunta central: “¿Según usted, cual es el problema más urgente a resolver en su comunidad?”.

- De ampliación: “¿Podría hacer una explicación más detallada?”.

- De esclarecimiento “¿A qué se refiere con menos corrupción?”.

3) Preguntas de estimación en una dirección (unipolar). El encuestado puede elegir una respuesta según un intervalo donde muestre que tan de acuerdo está con la pregunta. Ejemplo.

“¿Como definiría usted su empleo actual?”

Muy bueno\_

Bueno\_

Regular\_

Malo\_

Muy malo\_

4) Preguntas de estimación en dos direcciones (bipolar).

“¿En su opinión, ¿Cómo es el servicio de transporte público en la actualidad?”

Muy Bueno\_

Bueno\_

Malo\_

Muy Malo\_

Insuficiente\_

Inexistente\_

5) Preguntas basadas en respuestas tipo acuerdo-desacuerdo. Siguen la construcción de preguntas basadas en las escalas de Likert o en la elaboración de índices, conformados por uno o más indicadores.

Ejemplo

“Seleccione su grado de acuerdo o desacuerdo con respecto a la siguiente pregunta”

“Es necesario cambiar el sistema ejecutivo de gobierno presidencial, donde una sola persona toma decisiones importantes, a una junta de gobierno, donde 3 o más personas son los encargados de tomar



decisiones nacionales”

Muy de acuerdo\_

De acuerdo\_

Ni de acuerdo ni en desacuerdo\_

En desacuerdo\_

Muy en desacuerdo\_

6) Preguntas de graduación en rangos. Indican que percepción de importancia según una escala tiene el encuestado acerca de una pregunta donde su opinión. Ejemplo:

“Indique según la escala del 1 al 5, su opinión acerca de cuáles deben ser las prioridades que debe tener la nación en su presupuesto anual. Donde “1” indica la máxima prioridad, y 5 la mínima prioridad”

Educación\_                  Inversión social\_                  Nueva Constitución\_

Aumento a congresistas\_          Salud\_                  Seguridad\_

7) Preguntas con diferencias semánticas. Se consulta sobre la percepción del encuestado sobre un tema, institución o persona. Ejemplo.

“Por favor, marque con un visto en el espacio indicado, la cualidad o las cualidades que según usted mejor definen al gobierno Peruano”

Dictatorial\_\_                  Democrático\_\_

Pobre\_\_                  Rico\_\_

Peligroso\_\_                  Pacifista\_\_

Corrupto\_\_                  Transparente\_

No Corrupto\_\_

Control de calidad al cuestionario y a los encuestadores

El control de calidad sobre el cuestionario consiste en seleccionar dentro del conjunto de encuestados a un reducido grupo de ellos para constatar que se cumplieran los siguientes requisitos:

- a) Verificar que las preguntas realizadas por los encuestadores hayan sido correctamente comprendidas.
- b) Revisar las preguntas de la encuesta y descartar aquellas que tienden a crear incomodidad, rechazo o cohiben a los encuestados.
- c) Revisar las respuestas a preguntas abiertas que pudiesen ser replanteadas como preguntas cerradas.
- d) Eliminar aquellas preguntas cuyas respuestas pudiesen ser triviales o iguales a otras.

También se debe realizar un control de calidad sobre el desempeño de los encuestadores para asegurarse que.

- a) las encuestas efectivamente fueron ejecutadas.
- b) Si los encuestados cumplen con las características que lo identifican como miembros del grupo que es objeto de estudio.
- c) Si los encuestadores cumplieron con las indicaciones del investigador al momento de ejecutar la encuesta.

## **La escala**

La escala es un término que se utiliza para nombrar a aquellos instrumentos elaborados para medir variables no cognitivas (denominadas latentes o no manifiestas) y poder cuantificarlas. Las variables no cognitivas pueden ser opiniones, conductas, preferencias, actitudes, intereses, etc. Su característica principal es que los individuos responden a una escala de categorías ordenadas y graduadas donde escogen aquellas que más se acercan a su posición con respecto al tema que está siendo medido. La puntuación total es la suma de las puntuaciones asignadas a cada categoría escogidas por cada participante, por tanto, no existen respuestas correctas o incorrectas.

## **Tipos de escala**

Para Campbell, para las investigaciones cuantitativas existen dos tipos de escalas

### **Escala por intervalos**

En este tipo de escala se le asigna un valor numérico al nivel de presencia de la unidad medida establecida, en la respuesta obtenida por el encuestado. Por lo general se utiliza para conocer en que grado de acuerdo o de desacuerdo, de satisfacción o de insatisfacción está con respecto a una o varias afirmaciones. Al no existir la valoración cero, la diferencia entre los elementos no es constante, pero si permite conocer cuanto mayor o menor es la cantidad de variable en un elemento. La escala por intervalos necesita que la encuesta esté diseñada con valores tanto verbales como numéricos.

}La escala de Likert, la Net Promoter Score y la Matriz de escala bipolar (Escala de Osgood) son

Ejemplos de escalas por intervalos.

a)Ejemplo de escala de Likert

¿Está usted satisfecho con la atención brindada en Home & Market?

Extremadamente Satisfecho\_

Muy Satisfecho\_

Satisfecho\_

Poco Satisfecho\_

Insatisfecho\_

b)Ejemplo de Escala Net Promoter Score (recomendación del local a allegados).

Mide la probabilidad de que una persona recomiende un negocio, producto o servicio a sus allegados en una escala del 0 al 10.

Recomendaría usted a la tienda Home & Market?



Donde 0-6 = detractores.

7-8 = pasivos

9-10 = promotores

$NPS = \%promotores - \%detractores$ .

c)Matriz de Escala Bipolar (Escala de Osgood)

Este tipo de preguntas se organizan sobre una tabla, donde para conocer su opinión sobre un producto o servicio se utiliza una valoración graduada.

¿Qué le parecido el café de Café Central?

	Poco	Mucho	Neutro	Poco	mucho	
Suave						Cargado
Insípido						Sabroso
Poco aromático						Aromático

### Escala de razón

En las escalas de razón tienen como aspecto diferenciante de las escalas de intervalos en que sí poseen un valor de cero absoluto, lo que implica la ausencia total de la variable medida. Esta escala permite conocer si un elemento es igual o diferente a otro dentro de la muestra, si la cantidad de variable presente en un elemento es mayor o menor que la presente en otro, e incluso en que proporción mayor o menor. La diferencia entre elementos medidos es constante pudiendo decir si es dos, tres, cuatro, etc., veces mayor uno del otro.

Elemento a los que se le pudiese aplicar una escala de razón son; salarios, volúmenes de venta, etc.

Ejemplo: ¿Cuántos ceviches puede comer al día?

- 1-2
- 2-3
- 3-4
- 4-5
- Mas de 5

### El inventario

Es un instrumento específicamente estructurado para medir variables de personalidad. Al igual que la escala, las respuestas no son correctas o incorrectas, sino que reflejan la conformidad o inconformidad del encuestado con las preguntas formuladas.

Por lo general los inventarios se estructuran en pruebas estandarizadas o test y que han sido diseñadas por investigadores y centros de investigación para medir un gran número de variables. Cada una de ellas tiene su propia forma de ser codificadas, aplicadas e interpretadas. Una de las más importantes es las pruebas de tipo proyectivas.



## Pruebas proyectivas

Estas pruebas están diseñadas para que el encuestado responda a una cantidad determinada de preguntas o estímulos para luego ser analizadas e interpretadas. Entre ellas se encuentran las Pruebas de Rorschach y el Test de Percepción Temática. Este tipo de pruebas requieren un entrenamiento importante y un conocimiento elevado por parte del investigador que considera aplicarla.

Dado la amplitud de aptitudes que pueden ser evaluadas y la complejidad que encierran, es común que las pruebas de inventario se restrinjan a un número limitado de tópicos.

Existen múltiples test o inventarios que pueden ser aplicados para analizar la personalidad de la población objetivo. Un compendio de ellos puede obtenerse en el link:

[https://www2.javerianacali.edu.co/sites/ujc/files/catalogo\\_liip\\_2018.pdf](https://www2.javerianacali.edu.co/sites/ujc/files/catalogo_liip_2018.pdf)

# CAPÍTULO 5

## Tratamiento de variables (Análisis Descriptivo)

La recolección de datos en las investigaciones sociales cuantitativas constituye una etapa de mucha relevancia; esto se debe a que de ello dependerá la confiabilidad del análisis posterior de los datos recopilados. Esto es evidente pues si los datos fueron recopilados de forma errónea o deficiente, ello redundaría en interpretaciones sesgadas o falsas al momento de aplicar el análisis sobre ellos.

Teniendo entonces la certeza de una correcta recolección de datos, adecuada éstas a los objetivos que guían la investigación, la siguiente etapa, en el desarrollo de un proyecto de investigación cuantitativa, es el análisis de los datos recopilados.

Como se ha mencionado, los objetivos de la investigación son quienes guían el proceso de análisis de los datos; son quienes definen qué se debe medir, cuáles datos son relevantes, etc. Para análisis de los datos recolectados en una primera aproximación, se utiliza el análisis descriptivo como herramienta para establecer ciertos comportamientos de los datos como pueden ser la distribución de las variables en la población o muestra tomada como sujeto de estudio; cuántos individuos o unidades se agrupan en ciertas categorías establecidas para las variables por el investigador o presentes de forma natural en segmentos de la población o muestra; cuál es la magnitud de la sumatoria de los valores medios acumulados, y que tan alejados se encuentran de aquellos valores que fungen como referencia. Estos son algunos de los procedimientos que se podían llevar a cabo mediante el análisis descriptivo de los datos.

Cuando la investigación se hace sobre un segmento de la sociedad, el análisis descriptivo permite mostrar características presentes en ciertos grupos de individuos a través de la variable objeto de análisis. Por ejemplo, si la repetición de cierta edad es relevante entre el grupo analizado, ello podría permitir el análisis de ciertos comportamientos presentes en el conjunto de los individuos conformantes del estudio.

En este capítulo se busca presentar las principales técnicas estadísticas de análisis de datos de una manera introductoria, que permita sentar las bases para el uso de técnicas de análisis estadístico más complejas en un desarrollo posterior.

### **Tratamiento de variables**

Uno de los análisis básicos que se ejecutan sobre los datos recopilados es lo referente a la distribución de variables sobre el conjunto de la población estudiada, sean estos grupos seleccionados según categorías propuestas por el investigador; formados de manera natural en la población o en muestras presentes en la misma.

Existen varias formas de analizar la manera en que una variable se distribuye sobre una población, sobre categorías naturales o diseñadas, o sobre muestras pertenecientes a una población dada. Vamos a exponer

dos de ellas: la distribución de frecuencias (compuestas por distribución de frecuencia absoluta, relativa y acumulativa) y distribución de porcentajes. Estas podrían responder a objetivos de investigación como conocer la posición social de un conjunto de individuos, sus fuentes de ingreso, etc.

### Distribución de frecuencias absolutas

La frecuencia absoluta ( $n_i$ ) es la que contabiliza cuántas veces está presente un valor  $x_i$  de la variable  $x$  en la población o muestra seleccionada en el estudio. Veamos este tipo de distribución investigando el salario en dólares por hora de un grupo de 120 personas cuyos ingresos estuvieron dados por los siguientes valores.

2, 4, 5, 7, 12, 13, 15, 22, 32

Se puede apreciar que no existe ningún tipo de categorías en los datos recabados; sin embargo, según el diseño de una investigación hipotética, dentro de los objetivos se establece conocer los ingresos por hora de las personas según unas categorías determinadas por intervalos de valor, entonces se pueden crear estas categorías para satisfacer este requisito de la siguiente manera

1-4; 5-9; 10-14; 15-19; 20-24; 25-29; 30-34

Estos datos se pueden representar en una tabla como se muestra a continuación.

**Tabla 5. Representación de las frecuencias absolutas del estudio.**

N.º de Intervalos	Intervalos de ingresos por horas (dólares)	Frecuencias absolutas ( $n_i$ )
1	1-4	42
2	5-9	31
3	10-14	15
4	15-19	11
5	20.-24	8
6	25-29	7
7	25-29	4
8	30-34	2
	Total	<b>120</b>

Se puede observar en la distribución de la frecuencia de los datos que existe un número importante de la población (más de un tercio) que gana entre 1 y 4 dólares la hora de trabajo. La tabla mostrada también destaca que existe un intervalo donde se concentra la mayoría de los datos. Esto quiere decir que se está en



presencia de una distribución unimodal. Cuando existe varias distribuciones de frecuencia similares entonces hablamos de frecuencias multimodales.

### Distribución de frecuencias relativas

Si queremos conocer a cuanto equivale cada intervalo con respecto al total de la población o muestra, entonces debemos conocer su distribución de frecuencias relativas. La frecuencia relativa ( $h_i$ ) se define como el porcentaje que representa un intervalo perteneciente a la población o muestra dentro de esta. En otras palabras, el porcentaje de frecuencia absoluta con respecto al total de la población o muestra. Se representa mediante la siguiente formula.

$$h_i = \frac{n_i}{n} \times 100 \quad (3)$$

Siendo  $n$  el número total de individuos de la muestra o población. Entonces, la tabla anterior puede verse de la siguiente manera.

**Tabla 7. Ingresos por hora del total de la población estudiada donde se muestras la frecuencia relativa**

N.º de Intervalos	Intervalos de ingresos por horas (dólares)	Frecuencia relativa ( $h_i$ )
1	1-4	$(42/120) \times 100 = 35\%$
2	5-9	$(31/120) \times 100 = 25,8 \%$
3	10-14	$(15/120) \times 100 = 12,5\%$
4	15-19	$(11/120) \times 100 = 9,2\%$
5	20.-24	$(8/120) \times 100 = 6,7\%$
6	25-29	$(7/120) \times 100 = 5,8\%$
7	25-29	$(4/120) \times 100 = 3,3\%$
8	30-34	$(2/120) \times 100 = 1,7\%$
<b>Total</b>		<b>(120) 100%</b>

El valor total de los porcentajes debe ser igual o muy aproximado al valor total de los integrantes de la población o muestra. Es importante resaltar el proceso de redondeo al cual fue sometida cada intervalo de frecuencia, por ejemplo, el intervalo n.º 2 el resultado sin aplicación de redondeo fue de 25,833..., se aplicó el redondeo por aproximación y se dejó el valor 25,8. En este ejemplo existen tres casos donde el segundo



decimal no afecta al primer decimal en el momento hacer el redondeo y tres casos donde sí se cambia el valor del primer decimal, por lo que existe un equilibrio en la contención de errores por aproximación en los datos. Por convención cuando se ajusta por aproximación un intervalo, el próximo se deja sin redondear a fin de evitar introducir errores de aproximación.

Es importante no olvidar colocar el valor total de individuos que forman cada intervalo, como se observa en la tabla anterior. La importancia de esto radica en que, si solo se conocen valores relativos, el valor de error puede variar entre datos. Por ejemplo, si entre 10 personas 7 personas tienen el mismo salario representarían el 70% de total; si solo se presenta el porcentaje podría creerse que se trata de un número mucho mayor de individuos.

Otro aspecto importante es la presentación de los intervalos en donde se distribuirán los valores obtenidos de la población o muestra. En el ejemplo utilizado se crearon intervalos del tipo excluyente por valores, es decir 1-4, 5-9, o 10-14. Esto se hace con propósito de no solapar valores por error. Otras propuestas como construir intervalos 1-5, 5-10 o 10-15 pudiesen no clarificar que en el primer intervalo los valores de 5 no serán tomados en cuenta pues formarían parte del segundo, y así sucesivamente.

Para culminar esta sección, es importante que los datos presentados sean comparados con otros reactivos o estímulos sobre la misma población o muestra, que permitan establecer normas que puedan dar significados de importancia a la investigación y evitar resultados que se expresen como “la mayoría de los encuestados opinan que...” “un importante grupo dice...”. Por ejemplo, como el caso presentado en el capítulo 3, si se consulta la opinión pública sobre el servicio de aseo urbano, recomendable consultar su opinión sobre otros servicios públicos como la telefonía estatal o los servicios de salud.

### **Distribución de frecuencias acumuladas**

Es una medida que refleja cómo se produce la sumatoria de los valores obtenidos sobre una población o muestra y se denota como  $N_i$ . Es decir, como la frecuencia absoluta o relativa acumulada, es la suma de todas las frecuencias absolutas o relativas hasta cubrir toda la población o muestra. Las mismas pueden ordenarse de forma ascendente o descendente según los requerimientos de la investigación. Se calcula mediante la siguiente ecuación:

$$N_i = \sum_{k=1}^i n_k \quad (4)$$

Tomando el ejemplo que se ha utilizado, se tiene que para las frecuencias relativas acumuladas tendremos los siguientes resultados.

**Tabla 8. Frecuencias relativas y frecuencias relativas acumuladas del estudio.**

N.º de Intervalos	Intervalos de ingresos por horas (dólares)	Frecuencias relativas ( $h_i$ )	Frecuencia acumulada ( $N_i$ ) %
1	1-4	35%	35
2	5-9	25,8 %	35+25,8=60,8
3	10-14	12,5%	60,8+12,5= 73,3
4	15-19	9,2%	73,3+9,2=82,5
5	20.-24	6,7%	82,5+6,7=89,2
6	25-29	5,8%	89,2+5,8=95
7	25-29	3,3%	95+3,3=98,3
8	30-34	1,7%	98,3+1,7=100
<b>Total</b>		<b>(120) 100%</b>	<b>100%</b>

Para las frecuencias absolutas acumuladas se obtiene la siguiente tabla.

**Tabla 9. Frecuencias absolutas y frecuencias absolutas acumuladas presentes en el estudio.**

N.º de Intervalos	Intervalos de ingresos por horas (dólares)	Frecuencias absolutas ( $n_i$ )	Frecuencia acumulada ( $N_i$ )
1	1-4	42	42
2	5-9	31	42+31=73
3	10-14	15	73+15=88
4	15-19	11	88+11=99
5	20.-24	8	99+8=107
6	25-29	7	107+7=114
7	25-29	4	114+4=118
8	30-34	2	118+2=120
<b>Total</b>		<b>(120)</b>	<b>120</b>

### Presentación de resultados del tratamiento de variables

Al realizar en este capítulo el tratamiento correspondiente a las representaciones de las distribuciones de frecuencias mediante tablas se utilizó una de las representaciones de resultados que pueden construirse para



presentar resultados del tratamiento hecho a alguna variable determinada. Existe otro medio para representar el tratamiento estadístico de las variables y sus resultados, este es la representación mediante gráficas.

La ventaja principal de la representación gráfica de los datos recolectados y tratados estadísticamente estriba en la facilidad de visualizado de resultados y su interpretación. Utilizando conceptos como longitud, área y volúmenes mediante su representación geométrica sobre un sistema de coordenadas, es más sencillo comunicar resultados de investigación. también es común apoyarse en sombreados, rayados o coloreados en áreas de interés para la investigación; sin embargo, hay que resaltar que esta facilidad de comprensión mediante la consistencia y forma que brindan los gráficos, se logra en detrimento de la exactitud que proporcionan las tablas de datos antes vistas. Por tanto, es aconsejable que todo gráfico esté acompañado con las tablas fuente de información.

Las representaciones gráficas pueden ser de las siguientes formas:

\*Pictogramas. Ejemplo

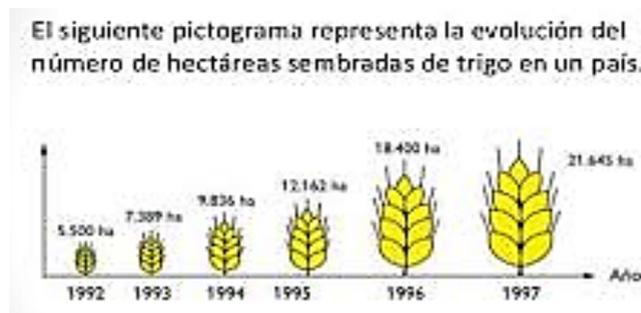


Figura 7. Pictogramas

\*Estereometría: incluyen elementos cúbicos, piramidales o prismáticos. Ejemplo

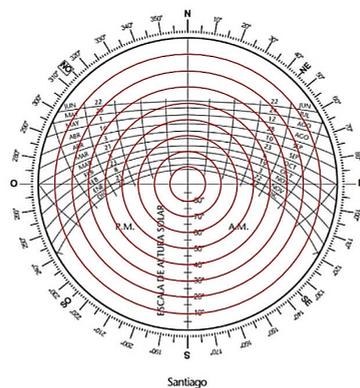


Figura 8. Estereometría

\*Cartografías: incluyen mapas estadísticos y cartodiagramas. Ejemplo

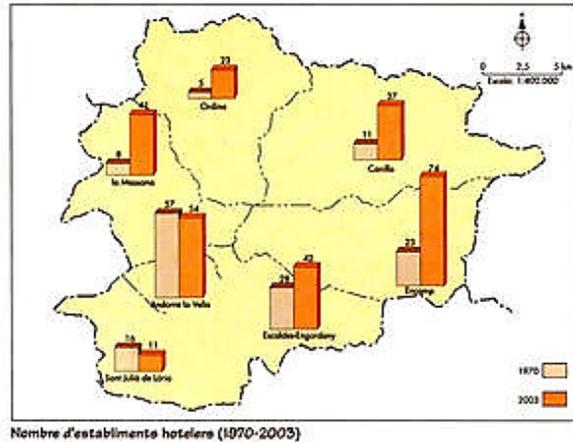


Figura 9. Cartografías

\*Diagramáticos: incluyen gráficas poligonales, de puntos, de barras, circulares, etc.

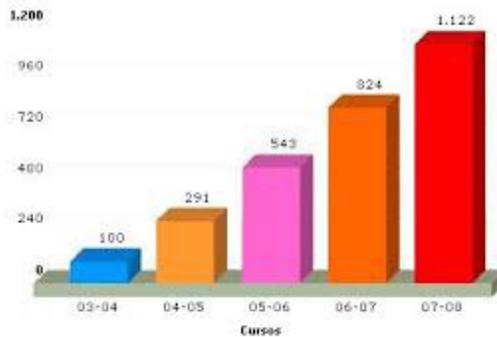


Figura 10. Diagramas

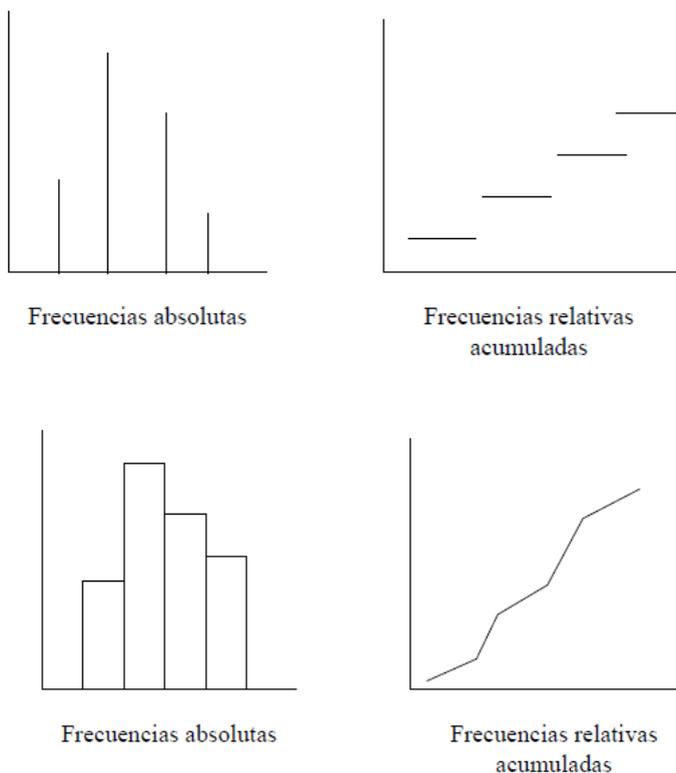
Los Diagramas de frecuencia son aquellos que representan a las variables cuando son discretas. Las frecuencias pueden ser, como se ha estudiado, absolutas o relativas acumuladas.

Cuando la(s) variable(s) a representar es(son) continua(s), se utilizan Histogramas de frecuencias, si utiliza(n) frecuencias absolutas. Si la frecuencia es relativa acumulada, entonces se utiliza una ojiva ascendente para representarla.

Si al histograma se le unen los puntos medios de la parte superior de cada rectángulo se obtiene el Po-

lígono de frecuencias.

Para resumir, las gráficas según la distribución de frecuencias pueden ser de la siguiente manera:



**Figura 11. Representación gráfica de resultados estadísticos de investigación básicos.**

En el análisis descriptivo de variables se utilizan ciertas medidas que permiten la caracterización descriptiva de un fenómeno que está siendo estudiado. Cuando se trabaja sobre poblaciones, las medidas son denominadas parámetros, y cuando se trabaja sobre muestras de una población, a las medidas se les denomina estadígrafos. Estas pueden estar presentes en distribuciones unidimensionales o pluridimensionales. En lo que respecta a los estadígrafos, estos se pueden clasificar según:

-Su posición. Estos son la media aritmética, la mediana, la moda, los cuartiles y los deciles.

-De concentración. Como el Índice de Gini.-

La dispersión. Estos son el rango, la varianza, la desviación estándar, el coeficiente de variación y las medidas de forma.

-La asimetría. Esto son el coeficiente de asimetría de Fisher, de Pearson y Bowley-Yule.

-De apuntamiento. De tipo leptocúrtica, mesocúrtica y platicúrtica.

A continuación, ampliaremos algunos de estos conceptos.

## Medidas de tendencia central

En los estudios de investigación social cuantitativa, es importante analizar el comportamiento estadístico de las variables. La creación de representaciones visuales que clarifiquen su distribución sobre la población o muestra estudiada, como son las tabulaciones (representación de datos mediante tablas) y las gráficas (representaciones sobre ejes coordenados) que son muy útiles para ese fin; sin embargo, además de la representación de los datos es importante su análisis mediante cálculos matemáticos que permitan extraer el comportamiento de las características presentes en la población o muestra sujeto de estudio. En un amplio número de estudios, los datos obtenidos de la población o muestra se agrupan entorno a un valor central, que permite obtener cierto valor característico o representativo del conjunto. A estos valores se les denomina medidas de tendencia central. Las medidas de tendencia central más importantes son: la moda, la mediana y la media aritmética. De ellas se hablará a continuación.

### Moda

En estadística se refiere al valor que más se repite en un intervalo de datos, es decir, es el atributo o valor que con mayor frecuencia se observa en un conjunto de datos. La moda se representa como  $M_0$ . Cuando en un determinado conjunto de datos se puede apreciar claramente la repetición una característica, atributo o valor por encima de otros, en ese caso se habla de una distribución unimodal. Si existen más de un atributo, característica o valor, que se repite dentro del conjunto de datos, se habla de distribución multimodal. Si no se aprecia ninguna tendencia entre los datos consultados se diría que dicho conjunto no tiene moda, o es amodal.

Para obtener la moda de un conjunto de datos, se debe crear la distribución de frecuencias y de ellos ubicar el atributo, valor o características con mayor frecuencia.

Si los datos se agruparon mediante intervalos de frecuencias, entonces se debe realizar lo siguiente:

1. Determinar el o los intervalos de mayor frecuencia absoluta  $n_i$ .
2. Calcular la moda o las modas utilizando la siguiente ecuación:

$$M_0 = l_i + \left( \frac{d_1}{d_1 + d_2} \right) \cdot a \quad (5)$$



Donde

$l_1$  : Límite inferior del intervalo con mayor frecuencia absoluta.

$d_1$ : Diferencia entre la mayor frecuencia absoluta y la anterior.

$d_2$ : Diferencia entre la mayor frecuencia absoluta y la siguiente.

$a$  : Amplitud del intervalo con mayor frecuencia absoluta.

Supongamos el siguiente ejemplo donde se está investigando cuánto invierte el comer una comunidad de Madrid. De la investigación se pudo construir la siguiente tabla.

**Tabla 10. Moda correspondiente a la comunidad de Madrid estudiada con respecto al tiempo que invierten en comer.**

Nº. de intervalos	Tiempo (min)	$n_i$	$f_i$	$N_i$	$F_i$
1	20-24	40	16%	40	16%
2	25-29	82	32.8%	122	48.8%
3	30-34	96	38.4%	218	87.2%
4	35-39	20	8%	238	95.2%
5	40-49	12	4.8	250	100%

1.Podemos observar que el intervalo donde existe una mayor frecuencia absoluta es el N.º 3.

2.Para conocer la moda identifiquemos los elementos que componen a la ecuación:

$$l_1 = 30.$$

$$d_1 = 96-82 = 14.$$

$$d_2 = 96-20 = 70.$$

$a = 34-29=5$ . (cada intervalo comprende tanto el valor inicial como el final, por ello la amplitud se mide desde el fin del intervalo anterior al próximo anterior al comienzo del intervalo con mayor frecuencia

absoluta)

sustituimos en la ecuación y obtenemos lo siguiente:

$$M_0 = 30 + \left( \frac{14}{70 + 14} \right) \cdot 5 \cong \mathbf{30.8} \text{ minutos}$$

Esto nos indica que las personas de la comunidad comen en un promedio de 30.8 minutos.

### **Mediana**

Es el valor central dentro de una distribución de valores ordenadas por lo general de menor a mayor, si el número de valores es impar, la mediana será el valor central; si el número de valores es par, la mediana será la semisuma de los dos valores centrales. Si los valores están dados por intervalos de frecuencia entonces se debe utilizar una ecuación que se mostrará más adelante en esta sección. En todo caso la mediana divide la cantidad de datos en dos partes: un 50% por debajo del valor de la mediana y otro 50% por encima de la mediana. La mediana se representa como Me. Veamos algunos ejemplos.

-Número impar de valores desordenados

20, 9, 10, 8, 13, 4, 6.

Ordenamos los valores y quedan: 4, 6, 8, 9, 10, 13, 20.

Ya ordenados podemos observar que la mediana es 9.

-Número par de valores desordenados

15, 22, 8, 11, 14, 18, 9, 24.

Ordenamos los valores y quedan: 8, 9, 11, 14, 15, 18, 22, 24.

Ya ordenados observamos que los valores centrales son 14 y 15. Se procede a realizar una semisuma  $(14+15) / 2 = 14.5$ .

-Cuando los valores están dados en intervalos de frecuencia se aplica el siguiente procedimiento:

1. Hallar a  $N/2$ .

2. Ubicar el intervalo cuya frecuencia absoluta acumulada  $N_i$  contiene a  $N/2$ .

3. Calcular la mediana mediante la ecuación



$$M_e = l_i + \left( \frac{\frac{N}{2} - N_{i-1}}{n_i} \right) \cdot a \quad (6)$$

$l_i$  : Límite inferior del intervalo que contiene  $N/2$ .

$N$  : Número total de datos de la población.

$N_{i-1}$  : Frecuencia absoluta acumulada antes del intervalo que contiene  $N/2$ .

$n_i$  : frecuencia absoluta del intervalo que contiene  $N/2$ .

$a$  : Amplitud del intervalo que contiene a  $N/2$ .

Utilizando como ejemplo la misma tabla del ejemplo de moda tenemos.

**Tabla 11. Frecuencias absolutas y relativas tanto por intervalos como acumuladas de la comunidad de estudio.**

Nº. de intervalos	Tiempo (min)	$n_i$	$f_i$	$N_i$	$F_i$
1	20-24	40	16%	40	16%
2	25-29	82	32.8%	122	48.8%
3	30-34	96	38.4%	218	87.2%
4	35-39	20	8%	238	95.2%
5	40-49	12	4.8	250	100%

Se sigue los pasos indicados anteriormente.

1.El número total de la población es de 250 personas. Entonces  $N/2= 250/2= 125$ .

2.Al observar los valores de la frecuencia absoluta podemos deducir que la persona 125 se encuentra en el intervalo N.º 3.

3. Los datos para el cálculo de la mediana serían:

$$l_i = 30.$$

$$N/2 = 125.$$

$$N_{i-1} = 122.$$

$$n_i = 96.$$

$$a = 5.$$

Entonces la mediana es

$$M_e = 30 + \left( \frac{125-122}{96} \right) \cdot 5 \cong 30.4 \text{ minutos}$$

Esto quiere decir que el 50% de las personas tardan menos de 30.4 minutos en comer y el otro 50% tarda más de 30.4 minutos en comer.

### Media aritmética

La media aritmética es la medida más empleada en los estudios de análisis estadístico y por tanto en la investigación social cuantitativa para identificar a un conglomerado mediante un único valor. Se calcula mediante la suma de todos los valores de los datos recabados, dividida entre el número total de la muestra. Si la variable de estudio se representa como  $X$ , la media aritmética se representaría como  $\bar{X}$ .

Como en los incisos anteriores, los datos se pueden presentar en diferentes formas.

Cuando los datos son pocos y no se han ordenado en intervalos. La ecuación que representa este caso es:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} \quad (7)$$

$\bar{X}$ : La media aritmética de la muestra.

$n$ : Total de datos que conforman a la muestra.

$x_i$ : Dato de la muestra.

$\sum_{i=1}^n x_i$  : sumatoria de todos los elementos de la muestra.

Si por ejemplo tenemos 8 corredores finalistas de los 100 metros planos y sus tiempos en segundos fueron los siguientes: 9, 8.8, 9.2, 10, 9.4, 9.8, 10.1, 9.6. podemos conocer la media aritmética de la siguiente manera:

$$\bar{X} = \frac{9 + 8.8 + 9.2 + 10 + 9.4 + 9.8 + 10.1 + 9.6}{8} \cong 9.49 \cong 9.5$$

En promedio los velocistas cubren los 100 metros en 9.5 segundos.

b) Cuando los datos se agrupan en una tabla de frecuencias sin intervalos, la media aritmética se calcula mediante la ecuación:

$$\bar{X} = \frac{\sum_{i=1}^m x_i \cdot n_i}{n} \quad (8)$$

Donde  $n_i$  representa a la frecuencia absoluta de cada valor de la variable.

Supongamos que la variable X sea la cantidad de vehículos que posee una familia en un conjunto habitacional H, representada en la tabla siguiente:

**Tabla 12. Número de vehículos por familia del Conjunto H.**

Número de autos $x_i$	Frecuencia $n_i$	$x_i \cdot n_i$
0	3	0x3=0
1	5	1x5=5
2	2	2x2=4
3	1	3x1=3
4	1	4x1=4
$n = \sum n_i = 12$		$\sum x_i \cdot n_i = 16$

Entonces

$$\bar{X} = \frac{16}{12} = 1,3$$

Significa que en promedio las familias del conjunto poseen un vehículo por familia.

c) Cuando los datos se agrupan en intervalos el cálculo de la media aritmética se realiza mediante la ecuación:

$$\bar{X} = \frac{\sum_{i=1}^m \dot{x}_i \cdot n_i}{n} \quad (9)$$

Donde  $(\dot{x}_i)$  es la marca de clase de cada intervalo. Entonces realizando el estudio sobre la tabla N.º 12 que se ha utilizado de manera recurrente tenemos.

Tabla 13. Media aritmética para intervalos de datos del caso de estudio.

Nº. de intervalo s	Tiempo (min)	$\dot{x}_i$	$n_i$	$\dot{x}_i \cdot n_i$
1	20-24	22.5	40	900
2	25-29	27.5	82	2255
3	30-34	32.5	96	3120
4	35-39	37.5	20	750
5	40-49	45	12	540
Total			$n = \sum n_i = 250$	$\sum x_i \cdot n_i = 7565$

La media aritmética sería

$$\bar{X} = \frac{\sum_{i=1}^m x_i \cdot n_i}{n} = \frac{7565}{250} = 30.26 \text{ minutos}$$

El resultado indica que en promedio las personas de esa comunidad de Madrid emplean un tiempo promedio de 30.26 minutos para comer.

### Medidas de dispersión

Las medidas de dispersión permiten conocer el nivel de homogeneidad o heterogeneidad de los datos recabados de una muestra o población, esto indica cuanta variabilidad existe en dichos datos. En este apartado se presentarán las medidas de dispersión más importantes, que son la varianza, la desviación estándar que son utilizadas en variables organizadas en intervalos o dependiendo de proporción, y el coeficiente de variación que se utilizan en variables ordinales o nominales.

#### La Varianza

Es la diferencia existente entre cada dato presente en la muestra con respecto a la media aritmética elevada al cuadrado. Este valor lo que nos indica es cuánta diferencia existe entre el valor consultado con respecto a promedio del conjunto. Para una muestra la varianza sería

$$S^2 = \frac{\sum(x_i - \bar{X})^2}{N} \quad (10)$$

Donde  $S^2$  es la varianza,  $x_i$  es un dato de la muestra,  $N$  es el número total de valores que conforman a la muestra y  $\bar{X}$  es la media aritmética.

Como es notorio, los datos fueron elevados al cuadrado y por ende las unidades también estarán elevadas al cuadrado, por lo que cualquier interpretación está imposibilitada. Por tanto, la varianza se utiliza como medida para comparar la dispersión existente entre dos o más variables, identificando dónde existe mayor varianza; sin embargo, la varianza es más útil como paso intermedio para conocer el cálculo de la desviación estándar de un conjunto de datos.

Veamos el siguiente ejemplo.

Se preguntó a un conjunto de 6 personas sobre el tiempo empleado en redes sociales, y estos fueron los resultados.

**Tabla 14. Distribución de horas en redes sociales del conjunto de personas estudiadas.**

N.º personas	Tiempo(horas)	Media de la muestra a	Desviación $(x_i - \bar{X})$	Desviación al cuadrado $(x_i - \bar{X})^2$
1	2.3	2.9	2.3-2.9= -0.6	0.36
2	3	2.9	3-2.9= 0.1	0.01
3	3.4	2.9	3.4-2.9= 0.5	0.25
4	3.5	2.9	3.5-2.9= 0.6	0.36
5	2.7	2.9	2.7- 2.9= -0.2	0.04
6	2,9	2.9	2.9-2.9= 0	0
			$\sum (x_i - \bar{X}) = 0.3$ $\approx 0$	$\sum (x_i - \bar{X})^2 =$ $1.02 \approx 1.0$

Entonces la varianza es

$$S^2 = \frac{\sum (x_i - \bar{X})^2}{N} = \frac{1}{6} = 0.1666 \approx 0.2 \text{ horas}^2$$

Observe que si deseamos interpretar la varianza diríamos que la variación entre el tiempo de exposición a redes sociales por parte de las personas objeto del estudio sería de 0.2 horas cuadradas, lo que no tiene sentido. Por ello es importante como transición al cálculo de la desviación estándar.

### Desviación estándar

Es la medida de dispersión más representativa de un conjunto de datos o valores. Se calcula aplicando la raíz cuadrada a la varianza y se denota matemáticamente como S cuando para una muestra.

Del ejemplo utilizado en la varianza tenemos que la desviación estándar sería

$$S = \sqrt{0.2} \cong 0.4 \text{ horas}$$

Se podría concluir que la exposición de las personas estudiadas varía en 0.4 horas alrededor de la media de 2,9 horas.

### **Coefficiente de variación**

No es posible comparar las desviaciones estándar de dos distribuciones diferentes de frecuencias de forma directa, pues dependen del valor de la media aritmética respectiva a cada una de ellas. Para ello se utiliza la siguiente ecuación.

$$CV = \frac{S}{\bar{X}} \cdot 100 \quad (11)$$

Si por ejemplo un grupo empresarial quiere invertir en proyectos de energía renovable por ello están estudiando dos proyectos eléctricos: el primero una central hidroeléctrica con utilidades promedio de 15 millones de dólares con una desviación de 5 millones de dólares. En el mismo año, una empresa eléctrica de energía solar tuvo una utilidad promedio de 7 millones de dólares con una desviación de 1 de dólares. ¿Cuál sería de ambos proyectos el más seguro para invertir?

Con estos datos podemos conocer el coeficiente de variación de la siguiente forma:

$$CV_{\text{empresa 1}} = \frac{5}{15} \cdot 100 = 33,3\%$$
$$CV_{\text{empresa 2}} = \frac{1}{7} \cdot 100 = 14.3\%$$

Realizando una comparación entre los coeficientes de variación de ambas empresas se puede notar que a pesar que la empresa de energía solar tiene menores beneficios posee una operación más estable por lo que generaría más confianza al grupo inversor.

### **Medidas de concentración**

Cuando se requiere conocer la concentración de los datos para una variable cuantitativa en una muestra, se emplean por lo general el Coeficiente de Gini o la Diferencia entre quintiles extremos en la distribución que posee la variable estudiada.

### **El Índice y Coeficiente de Gini**

El coeficiente de Gini es una herramienta que permite conocer el grado de desigualdad existente en una población con respecto a ciertos parámetros como la pobreza, el salario, el acceso a la educación, a la vivienda etc. Se construye en un rango que va del 0, que representa una perfecta igual, y 1 que representa desigualdad total (esto implica que un grupo reducido posee acceso a los bienes, educación, etc., y los demás

no reciben nada). El índice de Gini se obtiene multiplicando el coeficiente de por cien.

La ecuación que describe al coeficiente de Gini es la siguiente:

$$G = 1 - \sum (p_{i+1} - p_i)(q_{i+1} + q_i) \quad (12)$$

Donde  $p_i$  representa cada proporción de personas acumuladas en cada grupo investigado (podrían representar universitarios, obreros, empresarios, etc.) y  $q_i$  es la proporción acumulada de la participación de cada grupo en la variable de estudio (salario, pobreza, educación, etc.)

Supongamos en el siguiente ejemplo describe un caso sobre grado de igualdad en una población para el acceso a la vivienda.

**Tabla 15. Distribución de los datos de la muestra estudiada.**

<b>Ocupación</b>	<b>Vivienda propia (%)</b>	<b>Personas con vivienda propia (%)</b>
Estudiante	12.5	1.3
Bachiller	32.5	3.4
Técnico medio	20.6	5,2
Técnico superior	17.4	22.6
Graduado universitario	10,3	27.4
Post graduado	6,7	40.1
	100	100

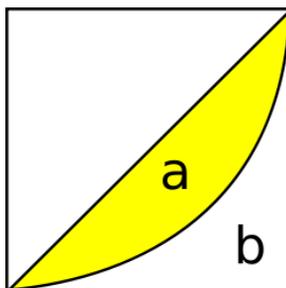
Convirtamos esta información en los parámetros necesarios.

**Tabla 16. Cálculo de los parámetros para conocer el Índice de Gini-**

$p_i$	$q_i$	$p_{i+1} - p_i$	$q_{i+1} + q_i$	$(p_{i+1} - p_i)(q_{i+1} + q_i)$
0.125	0.013	-	-	-
0.450	0.047	0.325	0.060	0.02
0.656	0.099	0.206	0.146	0.03
0.830	0.325	0.174	0.424	0.07
0.933	0.599	0.103	0.924	0.10
1	1	0.067	1.599	0.11
				0.33

El índice Gini sería  $1 - 0,33 = 0,67$ , lo que refleja un nivel de desigualdad importante.

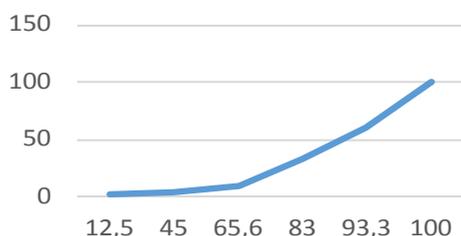
La representación gráfica del coeficiente de Gini es la denominada Curva de Lorenz donde el eje de las abscisas representa el porcentaje de la población y el eje de las ordenadas los valores de la variable estudiada cuya concentración es la que se desea conocer. Es frecuente añadir una recta diagonal del origen en 0 hasta el valor 1 y que expresa lo que se denomina distribución perfecta, que denota una igualdad perfecta. Entonces el área debajo de ella es la que indica el nivel de desigualdad. La curva de Lorenz se puede apreciar esquemáticamente a continuación.



**Figura 12. Representación de la curva de Lorenz**

Donde el coeficiente de Gini es  $a/(a+b)$ .

La curva de Lorenz para el estudio realizado queda de la siguiente manera



**Figura 13. Curva de Lorenz obtenida de los datos**

### Quintiles

Otro procedimiento para identificar si en un número de datos de una variable de estudio existe concentración es dividir en 5 partes el total de la población o lo que es lo mismo, dividir en 5 grupos donde cada uno de ellos corresponde con el 20% del total. A esto se llama dividir en quintiles. Luego con los valores obtenidos de la variable se calcula los totales de la variable en cada grupo. El nivel de concentración se encuentra comparando los valores ubicados en los quintiles de los extremos inferior y superior.

Por ejemplo, una variable vacunación en un conjunto de niños de ciertos colegios de una región según la siguiente distribución en los extremos.

**Tabla 17. Tabla del primer quintil y del último quintil de los colegios estudiados.**

Quintiles	Total, de vacunas colocadas en los niños	Total, de vacunas según los planteles
Primer quintil. Colegios del 1 al 6	400	13,3
Quinto quintil, Colegio de 21-26	1200	32.3
total		100 (11545)

Las cifras indican que los primeros colegios censados recibieron el 13.3% de las vacunas mientras que los últimos colegios recibieron el 32.3%

# CAPÍTULO 6

## Asociaciones y correlaciones

### Análisis inferencial

#### Introducción

En las investigaciones sociales es frecuente encontrar planteamientos de investigación que giran en torno al análisis de las asociaciones o correlaciones existentes entre dos o más variables, en donde se puedan apreciar interacciones interesantes y/o importantes para los objetivos buscados en el estudio.

Las asociaciones son aquellas relaciones que existen entre variables que no comparten características similares como pueden ser el género, la confesión religiosa, la región, etc. Además, otras relaciones que pueden ser consideradas como asociaciones son aquellas donde las variables son susceptibles a poder organizarse según un grado de incidencia como la satisfacción del cliente, el dominio de una tarea, etc.

Las correlaciones son aquellas relaciones que existen entre variables que fluctúan en un intervalo donde no existe un cero absoluto, como la temperatura, o que posean un carácter de proporcionalidad no nula como la edad, el peso, etc.

En general, en las asociaciones o correlaciones entre dos variables (menos las de carácter nominal) una de ellas tomará el papel de la variable dependiente y otra la de variable independiente.

Cuando se estudia la relación existente entre dos variables se toman en cuenta los siguientes aspectos que caracterizan dicha relación comprobaciones experimentales que prueben su validez y de esta manera pueda ser aceptado como parte constituyente de una teoría formal.

-Si existe realmente alguna relación y su grado, es decir, si se observa alguna variación en una variable cuando se produce una fluctuación o cambio en la otra, y si efectivamente existe la relación, el grado o coeficiente que expresa dicha relación (por lo general de una escala de referencia con valor máximo igual a 1).

-Orientación de la relación. Por lo general pueden ser de dos tipos. Una orientación directa, que indica que, si una variable fluctúa de forma incremental o decremental, la otra variable fluctúa de la misma forma. Y por otro lado está la orientación inversa, que significa que, si una variable se incrementa, la otra se decrementa, y viceversa.

-Funcionalidad de la relación. Este aspecto toma en cuenta la manera en que suceden los incrementos y los decrementos de las variables. Si los incrementos o decrementos poseen cierta uniformidad entre las variables, entonces su relación es lineal. Si luego de ciertos valores las diferencias incrementales aumentan o disminuyen de un dato a otro de forma considerable entre variables, entonces su relación es de tipo curvilíneo.

Cuando las investigaciones sociales poseen un carácter explicativo, el análisis dependerá del tipo de datos manejados, si son experimentales o no experimentales. Según el escenario, la investigación debe reflejar la covarianza existente entre las variables, en donde pueda observar la incidencia de relación temporal de una con respecto a la otra, antecediendo primero la variable independiente a la dependiente, y demostrando que dicha relación existe y no es producto de ninguna otra variable o evento que influya sobre ellas.

### Comparación de muestras

Para realizar cualquier proceso de investigación donde los datos obtenidos son probabilísticos es necesario pruebas que aclaren si las diferencias observadas entre muestras extraídas de un universo o población específica. Para resolver estos problemas de investigación se realizan test o pruebas de significancia estadística que tienen una secuencia de pasos para realizar cálculos y obtener resultados y de esa manera poder exponer conclusiones. El proceso es el siguiente:

1) Se formula una hipótesis nula y una hipótesis alterna. En la hipótesis nula se asume que no existen diferencias entre los valores de las muestras tomadas y en la hipótesis alterna se asume que sí existen diferencias y que son estadísticamente significativas.

2) Se elige una distribución de muestreo conocido que permita conocer la probabilidad que un valor perteneciente a una muestra pueda ocurrir dentro de su población. Si se toman muchas muestras con igual proporción dentro de una población esas muestras tenderán a tener una distribución normal. Un ejemplo lo podemos ver en la Fig. 14. para un estudio de estaturas en Argentina para una población femenina.

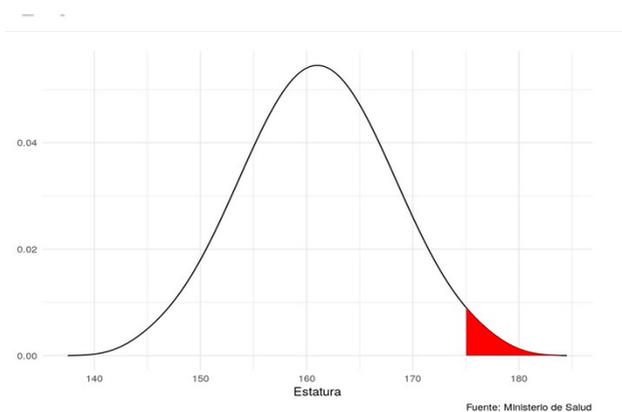


Figura 14. Estaturas de argentinas, la media es 1,75 cm.

3) Elegir un nivel de significancia. Las pruebas de significancia son parte de la teoría de probabilidades, y ayudan al investigador a saber si un valor se puede conseguir dentro de una distribución de datos aleatoria dentro del estudio ejecutado en una población. De esto se desprende el riesgo probabilístico de encontrar o no datos que respalden la hipótesis de la investigación. Entonces, el valor buscado convertido en una probabilidad estadística denominada sigma (llamada generalmente como estadística F o t) se expresa en un porcentaje (por lo general 5% del total de valores posibles. Aunque no es el único valor es uno de los más utilizados).

4) Cálculo de la estadística de la investigación. Se realizan mediante ecuaciones especiales que consisten en medir las diferencias observadas en las muestras mediante el error estándar de esa distribución según los datos recolectados en la investigación.

5) Verificación o descarte de la hipótesis nula. Cuando el valor estadístico z, t o F es mayor al supuesto en las tablas estadísticas de referencia se rechaza la hipótesis nula. Si por el contrario la comparación indica que el valor estadístico encontrado en la investigación es menor quiere decir que la hipótesis nula es correcta.

### Comparación entre dos proporciones

Supongamos que un estudio busca conocer en una parroquia urbana las probabilidades de que sus vecinos opten por emigrar. De una muestra de 90 personas de bajos recursos, el 15% piensa emigrar dentro de dos años. Otra muestra de 110 personas con recursos medios, el 25% piensa emigrar en los próximos dos años. Se quiere saber si hay una diferencia estadística significativa entre ambas muestras tomadas de la misma población.

a) Se procede entonces al cálculo estadístico para comprobar la premisa nula, que no existe diferencia entre muestras, o la hipótesis alterna, que existe una diferencia entre las muestras y ella es significativa estadísticamente.

b) Se asume un nivel de significancia del 5% (0.05).

c) Se determina el valor de z con la ecuación

$$z = \frac{p_1 - p_2}{\sqrt{P_u Q_u \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (13)$$

Donde  $p_1$  es la proporción o porcentaje mayor,  $p_2$  la proporción menor,  $n_1$  tamaño de la muestra de la proporción mayor,  $n_2$  tamaño de la muestra de la proporción menor,  $P_u$  es la estimación de la proporción mayor de la población y  $Q_u$  es el complemento y se calcula como  $Q_u = 1 - P_u$ .

d) Se procede al cálculo de  $P_u$

$$P_u = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{(110)(0,25) + (90)(0,15)}{110 + 90} = 0,21$$

e) Se calcula el complemento  $Q_u$

$$Q_u = 1 - P_u = 0,79$$

f) Ahora se calcula el valor de  $z$ .

$$z = \frac{0,25 - 0,15}{\sqrt{(0,21)(0,79) \left(\frac{1}{110} + \frac{1}{90}\right)}}$$
$$z = \frac{0,1}{0,06} = 1,67$$

g) Conclusión. Como el valor de 1,67 es mucho mayor a 1,96 que es el valor para una significancia de 5% en la tabla estadística correspondiente quiere decir que existe una diferencia entre las muestras extraídas de la población por lo que además no existe un error de muestreo. (Verificar en las tablas estadísticas para  $z$ ).

La comparación para dos medios aritmético se realiza de la misma manera, pero utilizando como referencia el valor estadístico  $t$  y la desviación estándar  $s$  de las submuestras del experimento. El valor  $t$  se calcula según la siguiente ecuación.

$$t = \frac{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}} \sqrt{\frac{n_1 + n_2}{n_1 n_2}}}{\quad} \quad (14)$$

Donde  $n$  son el tamaño de las submuestras y  $s$  las respectivas desviaciones estándar. Es importante que el investigador tome en cuenta los siguientes elementos.

a) Las variables deben ser medidas por intervalos o que sean proporcionales los cuales son necesarios para los cálculos de las medias aritméticas y de las varianzas.

b) Las mediciones que serán comparadas para cada submuestra deben ser independientes.

c) Las muestras deben ser probabilísticas.

d) Las varianzas de las muestras deben ser iguales o con diferencias pequeñas para evitar que se puedan establecer diferencias que no provengan de las medias aritméticas reales de las submuestras comparadas.

## Asociaciones

### Comparación entre porcentajes

Una de las técnicas más usadas en los estudios sociales, por su utilidad en la identificación de asociaciones entre variables, es la tabla de contingencia, que consiste en un ordenamiento por celdas donde se vacían los datos procedentes de la asociación de variables (cruce simultáneo) y que son procesados según método para luego ser tabulados (tabulación cruzada). Como las investigaciones sociales utilizan de forma extendida variables que se definen a nivel ordinal o nominal, la tabla de contingencia es particularmente útil.

Las tablas de contingencia se construyen en formatos de celdas  $n \times m$  o  $n \times n$  de filas y columnas. El caso más simple de estudio para los problemas lógicos del análisis es la tabla de contingencia  $2 \times 2$  o tabla bivariada. Bajo este marco se pueden realizar las comparaciones de porcentajes entre variables.

Como se indicó en la introducción del capítulo, en las investigaciones sociales las asociaciones y las correlaciones entre dos variables (relación bivariada) se estudian desde el punto de vista en la que una variable actúa como la variable independiente y la otra como variable dependiente. Esta postura sigue la regla de Zeisel de causa-efecto que de manera sucinta indica que, en un cuadro dimensional de dos factores o variables, uno de ellos se comporta como consecuencia de la distribución del otro factor o variable. El siguiente ejemplo ilustra este concepto mediante la comparación de proporciones o porcentajes en un caso de asociación de tipo lineal.

### Ejemplo

Se estudia el acceso a internet de hogares ubicados en una región del norte de Brasil que luego de tabulados los datos se obtiene la siguiente tabla de contingencia.

**Tabla 18. Acceso por internet por hogares de la región estudiada.**

Internet	Si	No	Total (%)
Urbano (%)	76,6 (25742)	23,3 (17538)	99,9 (43280)
Rural (%)	25,6 (7918)	64,3 (42640)	99,9 (50558)
<b>Total</b>	<b>33660</b>	<b>60178</b>	<b>99,9 (93838)</b>

Si se analizan los datos de manera horizontal, se puede observar que la ubicación del hogar actúa como la variable independiente y el acceso a internet actúa como la variable dependiente. Se puede apreciar que existe un mayor acceso a internet (63.7%) en las zonas urbanas que en las zonas rurales (23.5%) y el recíproco es cierto también, existen menos hogares sin internet en zonas urbanas (26.2%) que en las zonas rurales (76.4%) por lo que se concluye que existe una asociación entre la ubicación de un hogar y su posibilidad de obtener acceso a internet.

## Comparaciones Múltiples

En algunos estudios es importante conocer la cantidad de personas que conforman un hogar en una determinada región. En el caso del ejemplo es importante conocer el porcentaje de acceso a internet de los hogares con más de 5 personas y aquellos con menos de 5 personas. Entonces la tabla quedaría de la siguiente manera.

**Tabla 19. Acceso a internet de la región estudiada según la composición del hogar.**

Internet	N.º de personas		Total (%)
	Si	No	
Urbano (%)	≤ 5	76,6	99,9
	> 5	72,8	99,9
Rural (%)	≤ 5	28,6	99,9
	> 5	20,4	99,9

Si realiza la comparativa los porcentajes no varían mucho en las zonas urbanas con respecto al acceso a internet de hogares con un núcleo familiar mayor que los que tienen un núcleo familiar más reducido (un 3.8 %). En cambio, es notable la diferencia en aquellos hogares rurales con un núcleo familiar mayor que aquellos con un núcleo familiar más pequeño (una brecha de 9.2 %).

Sin embargo, la prueba más aceptada para identificar si existe asociación entre variables es la prueba o test de ji cuadrado ( $X^2$ ), la cual se expresa de la siguiente manera.

$$X^2 = \sum \frac{(f_o - f_e)^2}{f_e} \quad (15)$$

Donde  $f_o$  es la frecuencia observada,  $f_e$  la frecuencia esperada. Supongamos la siguiente tabla de contingencia recaba de un estudio sobre el efecto de una medicina A sobre un grupo de pacientes enfermos y no enfermos. Tenemos que:

**Tabla 20. Tabla de contingencia comparativa para las pruebas con la medicina A.**

<b>Paciente</b>	<b>Medicina A (Personas)</b>	<b>Placebo (Personas)</b>	<b>Total (Personas)</b>
Enfermo	30	40	70
No enfermo	90	80	170
<b>Total</b>	<b>120</b>	<b>120</b>	<b>240</b>

Se quiere comprobar si la medicina A causa algún efecto positivo sobre los pacientes enfermos donde el placebo es la medida de control del experimento. Esto quiere decir que se investiga si existe alguna asociación entre la medicina A y los pacientes. Entonces se plantean dos posibles hipótesis: 1) que la medicina A no causa ningún efecto y 2) que la medicina A causa algún efecto. Para verificar la primera hipótesis se necesitan realizar algunos cálculos. El primero es verificar el porcentaje de incidencia o frecuencia esperada para esta hipótesis la cual se calcula de los totales de las personas que se mantuvieron que enfermaron y las que no. El segundo cálculo tiene que ver con los valores esperados en cada tabulación de datos según la distribución de los porcentajes. Entonces.

**Tabla 21. Tabla de contingencia comparativa para las pruebas con la medicina A con porcentajes.**

<b>Paciente</b>	<b>Medicina A (Personas)</b>	<b>Placebo (Personas)</b>	<b>Total (Personas) (%)</b>
<b>Enfermo</b>	30	40	70 (29)
<b>No enfermo</b>	90	80	170 (71)
<b>Total</b>	<b>120</b>	<b>120</b>	<b>240 (100)</b>

Los cuales se calcularon como  $(70/240) * 100 = 29$ , para lo que enfermaron; y el resto (71) para los que no enfermaron. Para conocer los valores o frecuencias esperadas por cada tabulación entre las personas que enfermaron y las que no dependiendo si tomaron la medicina A o el placebo se procede de la siguiente manera. El 29% de 120 (para aquellos que tomaron la medicina A) es de 35, haciendo una aproximación para no utilizar decimales. Lo mismo se realizar para aquellos que tomaron el placebo y como son un grupo de 120 personas tenemos es mismo resultado obtenido anteriormente. La tabla queda de la siguiente manera.

Tabla 22. Tabla de contingencia comparativa para las pruebas con la medicina A con frecuencias esperadas.

Paciente	Medicina A (Personas)	Placebo (Personas)	Total (Personas) (%)
Enfermo	30	40	70 (29)
Frecuencia esperada	35	35	
No enfermo	90	80	170 (71)
Frecuencia esperada	85	85	
<b>Total</b>	<b>120</b>	<b>120</b>	<b>240 (100)</b>

Ahora podemos realizar la prueba de  $X^2$  el la sumatoria expresada en la ecuación W. su resultado sería

$$X^2 = 2,02 \text{ como valor crítico.}$$

Debemos conocer también el grado de libertad que se tiene para estos datos que se calcula como  $(\# \text{ de filas} - 1) * (\text{número de columnas} - 1)$ . Para este ejercicio el grado de libertad es igual a  $k=1$ . Debemos suponer un grado de significancia para estos datos y supongamos que requerimos un grado de significancia del 25% (0.25 en la tabla de valores de ji cuadrado). Según la tabla que se muestra a continuación en la Figura 23, se tiene que para  $k=1$  y una significancia del 25% el valor es de 1.32. como el valor está por encima del punto de referencia entonces existe asociación entre las variables estudiadas.

P = Probabilidad de encontrar un valor mayor o igual que el chi cuadrado tabulado, v = Grados de Libertad

w/p	0,001	0,0025	0,005	0,01	0,025	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5
1	10,8274	9,1404	7,8794	6,6349	5,0239	3,8415	2,7055	2,0722	1,6424	1,3233	1,0742	0,8735	0,7083	0,5707	0,4549
2	13,8150	11,9827	10,5965	9,2104	7,3778	5,9915	4,6052	3,7942	3,2189	2,7726	2,4079	2,0996	1,8326	1,5970	1,3963
3	16,2660	14,3202	12,8381	11,3449	9,3484	7,8147	6,2514	5,3179	4,6416	4,1083	3,6649	3,2831	2,9462	2,6430	2,3640
4	18,4662	16,4238	14,8602	13,2767	11,1433	9,4877	7,7794	6,7449	5,9886	5,3853	4,8784	4,4377	4,0446	3,6871	3,3567
5	20,5147	18,3854	16,7496	15,0863	12,8325	11,0705	9,2363	8,1152	7,2893	6,6257	6,0644	5,5731	5,1319	4,7278	4,3515
6	22,4575	20,2491	18,5475	16,8119	14,4494	12,5916	10,6446	9,4461	8,5581	7,8408	7,2311	6,6948	6,2108	5,7652	5,3481
7	24,3213	22,0402	20,2777	18,4783	16,0128	14,0671	12,0170	10,7479	9,8032	9,0371	8,3834	7,8061	7,2832	6,8000	6,3458
8	26,1239	23,7742	21,9549	20,0902	17,5345	15,5073	13,3616	12,0271	11,0301	10,2189	9,5245	8,9094	8,3505	7,8325	7,3441
9	27,8767	25,4625	23,5893	21,6660	19,0228	16,9199	14,6537	13,2880	12,2421	11,3857	10,6564	10,0060	9,4136	8,8632	8,3428
10	29,5879	27,1119	25,1881	23,2093	20,4832	18,3070	15,9872	14,5339	13,4420	12,5489	11,7887	11,0973	10,4732	9,8922	9,3418
11	31,2635	28,7291	26,7569	24,7250	21,9200	19,6752	17,2750	15,7671	14,6314	13,7007	12,8987	12,1836	11,5298	10,9199	10,3410
12	32,9092	30,3182	28,2997	26,2170	23,3367	21,0261	18,5493	16,9893	15,8120	14,8454	14,0111	13,2661	12,5838	11,9463	11,3403
13	34,5274	31,8830	29,8193	27,6582	24,7356	22,3620	19,8119	18,2020	16,9848	15,9839	15,1187	14,3461	13,6356	12,9717	12,3398
14	36,1239	33,4262	31,3194	29,1412	26,1189	23,6848	21,0641	19,4062	18,1808	17,1169	16,2221	15,4209	14,6553	13,9961	13,3393
15	37,6978	34,9494	32,8015	30,5780	27,4884	24,9958	22,3071	20,6030	19,3107	18,2451	17,3217	16,4940	15,7332	15,0197	14,3389
16	39,2518	36,4555	34,2671	31,9999	28,8453	26,2962	23,5418	21,7931	20,4651	19,3689	18,4179	17,5646	16,7795	16,0425	15,3385
17	40,7911	37,9462	35,7184	33,4087	30,1910	27,5871	24,7690	22,9770	21,6146	20,4587	19,5110	18,6330	17,8244	17,0646	16,3382
18	42,3119	39,4220	37,1664	34,8052	31,5264	28,8693	25,9894	24,1555	22,7895	21,6049	20,6014	19,6993	18,8679	18,0860	17,3379
19	43,8194	40,8847	38,5821	36,1908	32,8523	30,1435	27,2036	25,3289	23,9004	22,7178	21,6891	20,7638	19,9102	19,1069	18,3376
20	45,3142	42,3358	39,9969	37,5663	34,1696	31,4104	28,4120	26,4976	25,0375	23,8277	22,7745	21,8265	20,9514	20,1272	19,3374
21	46,7963	43,7749	41,4009	38,9322	35,4789	32,6706	29,6151	27,6620	26,1711	24,9348	23,8578	22,8876	21,9915	21,1470	20,3372
22	48,2676	45,2041	42,7957	40,2894	36,7807	33,9245	30,8133	28,8224	27,3015	26,0395	24,9390	23,9473	23,0307	22,1663	21,3370
23	49,7276	46,6231	44,1814	41,6383	38,0756	35,1725	32,0069	29,9792	28,4288	27,1413	26,0184	25,0655	24,0689	23,1822	22,3369
24	51,1790	48,0336	45,5584	42,9798	39,3641	36,4159	33,1992	31,1325	29,5833	28,2412	27,0960	26,0625	25,1064	24,2037	23,3367
25	52,6187	49,4351	46,9280	44,3140	40,6465	37,6525	34,3816	32,2825	30,6752	29,3388	28,1719	27,1183	26,1430	25,2218	24,3366
26	54,0511	50,8291	48,2898	45,6416	41,9231	38,8851	35,5632	33,4795	31,7946	30,4346	29,2463	28,1730	27,1789	26,2395	25,3365
27	55,4751	52,2152	49,6469	46,9628	43,1945	40,1133	36,7412	34,5796	32,9117	31,5284	30,3193	29,2265	28,2141	27,2569	26,3363
28	56,8918	53,5939	50,9936	48,2782	44,4608	41,3372	37,9159	35,7150	34,0266	32,6295	31,3909	30,2791	29,2486	28,2740	27,3362
29	58,3006	54,9662	52,3355	49,5878	45,7223	42,5569	39,0875	36,8598	35,1394	33,7109	32,4612	31,3305	30,2825	29,2908	28,3361

Figura 23. Tabla de distribución ji cuadrado para diferentes grados de libertad

## **Análisis de varianza**

La varianza, como se indicó en el capítulo anterior busca conocer cuál es el grado de dispersión de los datos recolectados en una investigación con respecto a la media aritmética del experimento. A mayor dispersión de los datos, menos representatividad tendrá la media aritmética en la investigación.

En algunas investigaciones sociales, el investigador necesita conocer dentro de una población la manera en que se distribuyen ciertos subgrupos o ciertas categorías, para ello requiere conocer tanto los valores que representan a dichos subgrupos como también las medias aritméticas correspondientes a cada uno de ellos. Por lo general el motivo es comparar las medias aritméticas mediante un cálculo estadístico y permite sacar conclusiones significativas de esa población. El uso de una técnica estadística para obtener un resultado que permita realizar conjeturas o identificar conclusiones sobre los valores manipulados se le denomina análisis de varianza. Dicho análisis puede ser de comparación simple si se estudia una sola variable independiente escogida por el investigador o puede ser de comparación múltiple si involucra dos variables. Se iniciará por el análisis de comparación simple.

### **Análisis de varianza simple**

En este análisis se trabaja con una sola variable independiente identificada por el investigador y de cuyas categorías se obtendrán los subgrupos que serán objeto de análisis, donde se realizarán las comparaciones de las medias aritméticas de cada subgrupo y corroborar si existen asociaciones significativas entre sus datos. Para ello se requiere conocer las medias aritméticas de cada subgrupo (sean iguales o no), subgrupos o muestras independientes y que cumplan con la condición de homocedasticidad, que es la igualdad de varianza o de desviación estándar para cada subgrupo o muestra.

Se procede como se recordará en el apartado de comparación de proporciones: 1) se establece una hipótesis nula y una hipótesis alterna, 2) cálculo de medias aritméticas verificando si existe igualdad o no, 3) utilización de la tabla de distribución F de Fisher, 4) elección del nivel de significancia, 5) identificación de los valores correspondientes al cálculo de la estadística F, 6) comparación del resultado con el valor de la tabla según el grado de libertad obtenido. Todo este procedimiento nos lleva a la verificación o rechazo de la hipótesis nula.

El procedimiento de análisis de la varianza comienza calculando la suma de las desviaciones con respecto a las medias aritméticas correspondientes a cada muestra o subgrupo de una población. Este cálculo está constituido por dos secciones: la primera, que realiza la suma de cuadrados de las diferencias entre las medias de cada subgrupo con su media aritmética. A esto se le llama Suma de Cuadrados Dentro (SCD). La otra sección, que resulta de la suma de cuadrados entre las diferencias de las medias aritméticas de cada subgrupo con respecto a la media aritmética del total de las medidas y a la que se denomina Suma de Cuadrados Entre (SCE). Con estos valores se puede conocer el valor de la varianza de cada dato dentro de la población estudiada y de la cual se identificaron los subgrupos o muestras, y con el uso de los grados de libertad ya mencionados se identifica el valor estadístico en la tabla F. observemos el siguiente ejemplo.

### **Ejemplo**

Supongamos que el Estado quiere evitar la migración de los campesinos de una zona agrícola afectada por las bajas tazas de productividad de la tierra. Promoviendo para ello el empleo de fertilizantes desarrollados en tres universidades del país, el Estado quiere asegurarse de usar el mejor fertilizante de los tres

disponibles, pero se quiere asegurar que existan diferencias de calidad entre ellos. Por tanto, le provee a los campesinos de muestras de los tres fertilizantes para cosechar en 4 parcelas de iguales características un mismo producto. Al final se realiza una prueba de la productividad alcanzada con una tabulación basada en 10 como máxima productividad para cada fertilizante y una significancia del 5% suponiendo que los tres fertilizantes son idénticos en efectos como hipótesis nula. Obteniéndose los siguientes datos.

**Tabla 23. Tabla correspondiente al uso de los fertilizantes**

Fertilizante 1		Fertilizante 2		Fertilizante 3	
$X_1$	$X_1^2$	$X_2$	$X_2^2$	$X_3$	$X_3^2$
6	36	3	9	10	100
7	49	4	16	8	64
8	64	5	25	6	36
9	81	7	49	9	81
$\Sigma$ 30	230	19	99	33	281

Procedimiento

1) Cálculo de la sumatoria total de cuadrados mediante la siguiente ecuación, donde N es el número total de datos recolectados.

$$TC = \sum X^2 - \frac{\sum(X)^2}{N} \quad (16)$$

Sustituyendo los datos se obtiene

$$STC = (230 + 99 + 281) - \frac{(30 + 19 + 33)^2}{12}$$

$$STC = 610 - 560$$

$$STC = 50$$

2) Se calcula la sumatoria de cuadrados dentro para cada uno de los grupos, utilizando la ecuación anterior



$$\text{Grupo 1} = (230) - \frac{(30)^2}{4}$$

$$\text{Grupo 1} = 230 - 225$$

$$\text{Grupo 1} = 5$$

$$\text{Grupo 2} = (99) - \frac{(19)^2}{4}$$

$$\text{Grupo 2} = 99 - 90$$

$$\text{Grupo 2} = 9$$

$$\text{Grupo 3} = (281) - \frac{(33)^2}{4}$$

$$\text{Grupo 3} = 281 - 272$$

$$\text{Grupo 3} = 9$$

La sumatoria total de cuadrados dentro sería igual a  $SCD = 5 + 9 + 9 = 23$

3) Se calcula la sumatoria de cuadrados entre (SCE), que resulta de la diferencia existente entre la suma total de cuadrados y el total de la suma de cuadrados dentro (SCD). El cálculo sería

$$SCE = STC - SCD$$

$$SCE = 50 - 23$$

(17)

$$SCE = 27$$

4) Los grados de libertad de la suma total (GLST) se calculan con el número total de datos menos uno es decir Grados de Libertad Total =  $N - 1$  donde  $N$  es el total de mediciones

$$GLST = 12 - 1$$

$$GLST = 11$$

Para calcular los grados de libertad de la suma de cuadrados dentro (GLSCD) es igual a  $k(n-1)$ , donde  $k$  es el número de grupos y  $n$  el tamaño de cada grupo. Tenemos entonces que

$$GLSCD = 3(4 - 1)$$

$$GLSCD = 9$$

Para calcular el grado de libertad de la suma de cuadrados entre (GLSCE) se obtiene mediante  $k-1$

$$GLSCE = 3 - 1$$

$$GLSCE = 2$$

Una vez obtenidos todos los valores requeridos, se construye una tabla de análisis de varianza.

Tabla 24. Análisis de varianza

Variación	Suma de cuadrados	Grados de libertad	Estimación o media	Factor de comparación F
Entre grupos	27	2	13.5	$\frac{13,5}{2.6} = 5.19$
Dentro de los grupos	23	9	2,6	
Total	50	11		

5) Conclusión. El valor obtenido de F se compara con la tabla de variación F de Fisher para una muestra aleatoria de sumatorias “entre” y “dentro” según varios grados de libertad y con significancia de 5 % como se observa en la Fig. 16. Para los grados de libertad del numerador (2) y el denominador (9) podemos ver que 5,19 es superior a 4,26, lo que indica que la hipótesis nula es incorrecta, por tanto, existen diferencias en los cultivos de diferentes fertilizantes creados en las universidades. Para conocer cuál de ellos es el mejor se deben aplicar técnicas adicionales como las de Tukey o Scheffe.

g.d.	Grados de libertad del Numerador														g.d.	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14		15
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,0	243,9	244,7	245,4	245,9	1
2	18,513	19,000	19,164	19,247	19,296	19,330	19,353	19,371	19,385	19,396	19,405	19,413	19,419	19,424	19,429	2
3	10,128	9,552	9,277	9,117	9,013	8,941	8,887	8,845	8,812	8,786	8,763	8,745	8,729	8,715	8,703	3
4	7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041	5,999	5,964	5,936	5,912	5,891	5,873	5,858	4
5	6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818	4,772	4,735	4,704	4,678	4,655	4,636	4,619	5
6	5,987	5,143	4,757	4,534	4,387	4,284	4,207	4,147	4,099	4,060	4,027	4,000	3,976	3,956	3,938	6
7	5,591	4,737	4,347	4,120	3,972	3,866	3,787	3,726	3,677	3,637	3,603	3,575	3,550	3,529	3,511	7
8	5,318	4,459	4,066	3,838	3,687	3,581	3,500	3,438	3,388	3,347	3,313	3,284	3,259	3,237	3,218	8
9	5,117	4,256	3,863	3,633	3,482	3,374	3,293	3,230	3,179	3,137	3,102	3,073	3,048	3,025	3,006	9
10	4,965	4,103	3,708	3,478	3,326	3,217	3,135	3,072	3,020	2,978	2,943	2,913	2,887	2,865	2,845	10
11	4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948	2,896	2,854	2,818	2,788	2,761	2,739	2,719	11
12	4,747	3,885	3,490	3,259	3,106	2,996	2,913	2,849	2,796	2,753	2,717	2,687	2,660	2,637	2,617	12
13	4,667	3,806	3,411	3,179	3,025	2,915	2,832	2,767	2,714	2,671	2,635	2,604	2,577	2,554	2,533	13
14	4,600	3,739	3,344	3,112	2,958	2,848	2,764	2,699	2,646	2,602	2,565	2,534	2,507	2,484	2,463	14
15	4,543	3,682	3,287	3,056	2,901	2,790	2,707	2,641	2,588	2,544	2,507	2,475	2,448	2,424	2,403	15
16	4,494	3,634	3,239	3,007	2,852	2,741	2,657	2,591	2,538	2,494	2,456	2,425	2,397	2,373	2,352	16
17	4,451	3,592	3,197	2,965	2,810	2,699	2,614	2,548	2,494	2,450	2,413	2,381	2,353	2,329	2,308	17
18	4,414	3,555	3,160	2,928	2,773	2,661	2,577	2,510	2,456	2,412	2,374	2,342	2,314	2,290	2,269	18
19	4,381	3,522	3,127	2,895	2,740	2,628	2,544	2,477	2,423	2,378	2,340	2,308	2,280	2,256	2,234	19
20	4,351	3,493	3,098	2,866	2,711	2,599	2,514	2,447	2,393	2,348	2,310	2,278	2,250	2,225	2,203	20
21	4,325	3,467	3,072	2,840	2,685	2,573	2,488	2,420	2,366	2,321	2,283	2,250	2,222	2,197	2,176	21
22	4,301	3,443	3,048	2,817	2,661	2,549	2,464	2,397	2,342	2,297	2,259	2,226	2,198	2,173	2,151	22
23	4,279	3,422	3,028	2,796	2,640	2,528	2,442	2,375	2,320	2,275	2,236	2,204	2,175	2,150	2,128	23
24	4,260	3,403	3,009	2,776	2,621	2,508	2,423	2,355	2,300	2,255	2,216	2,183	2,155	2,130	2,108	24
25	4,242	3,385	2,991	2,759	2,603	2,490	2,405	2,337	2,282	2,236	2,198	2,165	2,136	2,111	2,089	25
26	4,225	3,369	2,975	2,743	2,587	2,474	2,388	2,321	2,265	2,220	2,181	2,148	2,119	2,094	2,072	26
27	4,210	3,354	2,960	2,728	2,572	2,459	2,373	2,305	2,250	2,204	2,166	2,132	2,103	2,078	2,056	27
28	4,196	3,340	2,947	2,714	2,558	2,445	2,359	2,291	2,236	2,190	2,151	2,118	2,089	2,064	2,041	28
29	4,183	3,328	2,934	2,701	2,545	2,432	2,346	2,278	2,223	2,177	2,138	2,104	2,075	2,050	2,027	29
30	4,171	3,316	2,922	2,690	2,534	2,421	2,334	2,266	2,211	2,165	2,126	2,092	2,063	2,037	2,015	30
31	4,160	3,305	2,911	2,679	2,523	2,409	2,323	2,255	2,199	2,153	2,114	2,080	2,051	2,026	2,003	31
32	4,149	3,295	2,901	2,668	2,512	2,399	2,313	2,244	2,188	2,142	2,103	2,070	2,040	2,015	1,992	32
33	4,139	3,285	2,892	2,659	2,503	2,389	2,303	2,235	2,179	2,133	2,093	2,060	2,030	2,004	1,982	33
34	4,130	3,276	2,883	2,650	2,494	2,380	2,294	2,225	2,170	2,123	2,084	2,050	2,021	1,995	1,972	34
35	4,121	3,267	2,874	2,641	2,485	2,372	2,285	2,217	2,161	2,114	2,075	2,041	2,012	1,986	1,963	35
40	4,085	3,232	2,839	2,606	2,449	2,336	2,249	2,180	2,124	2,077	2,038	2,003	1,974	1,948	1,924	40
60	4,001	3,150	2,758	2,525	2,368	2,254	2,167	2,097	2,040	1,993	1,952	1,917	1,887	1,860	1,836	60
80	3,960	3,111	2,719	2,486	2,329	2,214	2,126	2,056	1,999	1,951	1,910	1,875	1,845	1,817	1,793	80
90	3,947	3,098	2,706	2,473	2,316	2,201	2,113	2,043	1,986	1,938	1,897	1,861	1,830	1,803	1,779	90
100	3,936	3,087	2,696	2,463	2,305	2,191	2,103	2,032	1,975	1,927	1,886	1,850	1,819	1,792	1,768	100
120	3,920	3,072	2,680	2,447	2,290	2,175	2,087	2,016	1,959	1,910	1,869	1,834	1,803	1,775	1,750	120
inf.	3,841	2,996	2,605	2,372	2,214	2,099	2,010	1,938	1,880	1,831	1,789	1,752	1,720	1,692	1,666	inf.

Figura 16. Tabla de distribución de Fisher para índice de veracidad del 5%.

### Análisis múltiple de varianza

Existen estudios que requieren estudiar cómo interactúa más de una variable independiente dentro de una población. En el apartado anterior observamos cómo se trata un estudio inferencial cuya varianza depende una sola variable independiente, que fue el uso de tres tipos de fertilizantes para un mismo tipo de cultivo en una zona agrícola específica. Supóngase que ahora aparte del fertilizante se añade un tipo de abono que puede ser orgánico o inorgánico. Entonces existen tres tipos de fertilizantes y dos tipos de abono por lo que se tendrán 6 tipos de cultivos y por tanto 6 posibles producciones diferentes en calidad. Podría ser uno de los objetivos de la investigación el conocer cuál sería la combinación de fertilizante y abono que generará una mayor producción, si existe tal combinación, o si la producción es indiferente de la combinación.

Para poder conocer qué resultado arrojan este posible estudio y a cuáles conclusiones se podrían llegar,

se requiere de un análisis de varianza múltiple, que consiste en un estudio similar al mostrado en el análisis de varianza simple, donde se busca conocer el valor estadístico F de manera similar a partir de los datos otorgados por el problema a estudiar.

### **Coefficientes de correlación para variable ordinales**

Hasta el momento se ha estudiado mediante análisis de varianza si existen asociaciones entre variables, pero no la magnitud de dicha asociación. Para conocer el grado de asociación entre variables se requiere del uso de los coeficientes de correlación. A continuación, serán mostrados alguno de ellos.

Una variable ordinal es aquella que puede ser categorizada según alguna relación de jerarquía, teniendo así un lugar que no puede ser alterado por otros factores. Por ejemplo cuando se trata la variable edad, es posible categorizarla en niños, jóvenes, adultos, ancianos. Ninguna de estas categorías se puede alterar de su orden de jerarquía, ya que la variable no puede ser adulta antes que niño, tampoco puede ser anciano antes que joven.

Partiendo de esta premisa es posible verificar la existencia de una asociación del comportamiento de las variables, de una respecto de la otra. Así, resulta posible verificar la frecuencia de una variable ordinal A, con otra variable ordinal B, evaluando la posible asociación directa, también conocida como asociación positiva, cuyo valor debe ser cerca de 1. Cuando los cambios que presenta una variable no influyen en la otra variable, entonces se dice que existe una independencia de una respecto de la otra, cuyo valor estará cercano a cero.

### **Método de pares de valores**

Para llevar a cabo este proceso de observación es necesario que se evalúen los posibles pares de variables, analizando el comportamiento de los valores de cada variable. Resulta necesario reconocer el total de pares de valores (T), tomando en cuenta que estos no pueden estar repetidos, y teniendo en cuenta el total de casos N. La ecuación (18) permite determinar el valor T de pares;

$$T = \frac{N(N - 1)}{2} \quad (18)$$

Si por ejemplo N es 5 entonces T será 10, por lo tanto habrá 10 combinaciones de pares de la variable, donde no se permiten repeticiones.

Por ejemplo si se tiene N=5 (a, b, c, d, e) tendremos que T será: ab, ac, ad, ae, bc, bd, be, cd, ce, de.

Así pues es posible considerar lo siguiente:

Pares semejantes ( $N_s$ ): Son aquellos pares que guardan una semejanza de orden en cada variable.

Pares no semejantes ( $N_{ns}$ ): son aquellos pares cuyo orden es contrario en cada variable.



Pares encontrados en la variable independiente: Se trata de aquellos pares que coinciden en la misma posición, con el mismo valor ordinal, pero únicamente en la variable independiente.

Pares encontrados en la variable dependiente: Consiste en aquellos pares que coinciden únicamente en la variable dependiente, con la misma posición y el mismo valor ordinal.

Pares en ambas variables: es cuando un par además de coincidir en la variable dependiente, coincide en la variable independiente. Este tipo de pares se calcula con la ecuación (19).

$$T_{xy} = \sum \frac{f(f-1)}{2} \quad (19)$$

### **Coefficiente tau-a de Kendall**

Está definido por la razón entre la diferencia de pares semejantes y no semejantes, con la cantidad de pares totales, como muestra la ecuación (20).

$$\tau_a = \frac{N_s - N_d}{T} \quad (20)$$

El valor del coeficiente puede variar entre 1 y -1, donde la asociación puede determinar el signo. Cuando el coeficiente se encuentra en cero, se dice que existe una independencia estadística. El coeficiente es simétrico, por lo que el resultado no se afecta con la variable dependiente o independiente. Sin embargo, es posible que cuando los valores centrales no tengan frecuencia cero, el coeficiente no alcance asociación perfecta.

### **Coefficiente Q (de Yule)**

Se utiliza para tablas de correlación 2 x 2 para variables dicotómicas, que pueden ser ordinales o nominales. La ecuación que la describe según el orden de valores en la tabla es:

$$Q = \frac{ad - bc}{ad + bc} \quad (21)$$

Si existe una de las celdas con un valor igual a cero los límites de valores que adquiere Q son  $\pm 1$ .

Ejemplo. En un universo de encuestados se obtuvieron los siguientes datos de personas con hábitos alimenticios saludables y no saludables.

**Tabla 26. Hábitos alimenticios de la población encuestada.**

Hábitos alimenticios	Hombres	Mujeres
Saludables	90	220
No saludables	230	120

$$Q = \frac{(90)(120) - (220)(230)}{(90)(120) + (220)(230)} = -0.65$$

Lo que se puede interpretar como una tendencia bastante marcada por parte de las mujeres en preferir hábitos alimenticios saludables en comparación con los hombres. Si todas las mujeres y los hombres marcaran solo una preferencia, el coeficiente sería igual a 1.

### **Coeficiente $\phi$ (fi)**

Es otra medida para variables dicotómicas de tipo ordinal, nominal o a intervalos. Los valores a, b, c y d son los mismos que en el coeficiente Q. Su ecuación es

$$\phi = \frac{ad - bc}{\sqrt{(a + b)(c + d)(a + c)(d + b)}} \quad (22)$$

Si sustituimos en la ecuación los datos obtenidos anteriormente, obtenemos Que  $\phi = -0.39$ . como se observa el valor disminuye considerablemente al obtenido por el coeficiente Q; esto se debe a que este coeficiente toma en cuenta los marginales entre variable, lo que a mayor discrepancia entre ellos menor es el valor que asume  $\phi$ .

### **Coeficiente gamma de Goodman y Kruskal**

Este coeficiente se encarga de corregir el problema de la frecuencias centrales. Es también considerado como un valor simétrico, sin embargo en cualquier circunstancia puede alcanzar el asociación perfecta, según la ecuación:

$$\gamma = \frac{N_s - N_d}{N_s + N_d} \quad (23)$$



## Coefficiente rho de Spearman

Este coeficiente es sumamente utilizado en tratamientos estadísticos. Se emplea para cuando la variable dependiente tiene atributos ordenados, en dos categorías de una variable independiente, que además pudiese ser nominal. Este coeficiente se puede calcular con algún procedimiento previo que exija el ordenamiento.

Ejemplo: Se le pide a un grupo de personas que califiquen de 1 a 10 el menú de un restaurante, donde 1 es la de mejor aceptación y 10 la de menor gusto por los comensales.

**Tabla 27. Datos recogidos del estudio ejemplo.**

Menú	Cocineros		Comensales		$d$	$d^2$
	Media	Orden	Media	Orden		
Plato 1	1,34	1	2,33	1	0	0
Plato 2	6,58	5	8,44	7	-2	4
Plato 3	5,43	3	6,70	5	-2	4
Plato 4	7,10	6	7,94	6	0	0
Plato 5	7,11	7	4,58	4	3	9
Plato 6	6,24	4	3,25	2	2	4
Plato 7	5,12	2	4,11	3	-1	1
				$\Sigma$	0	22

Se evaluaron a un cierto grupo de personas, y sus respuestas se agruparon según la selección, a estos valores se le calculó la media y se ordenaron los datos según la media obtenida. Luego se calculó la diferencia  $d$  entre los grupos de estudio. Luego se aplica la ecuación (24) que corresponde a rho de Spearman.

$$\gamma = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \quad (24)$$

Para el ejemplo mostrado el valor de  $\gamma$  está dado por 0,61, que corresponde a un valor intermedio entre el gusto de los comensales y el de los cocineros. En el caso de que ambos grupos tuviesen el mismo gusto, el resultado habría sido 1, y en el caso en que ambos grupos tuviesen opiniones totalmente opuestas, el resultado habría sido -1.

## Análisis de regresión

El análisis de regresión es una técnica utilizada en los procesos estadísticos, que permite analizar el comportamiento de una variable, considerada dependiente, con relación a una o varias variables independientes. Este análisis se realiza variando o modificando el valor de la variable dependiente y dejando fija las variables independientes. De esta manera es posible estimar el valor promedio de la variable dependiente, cuando las variables independientes mantienen sus valores estables o fijos. Esta estimación del valor promedio se realiza para encontrar la función de regresión, que permita caracterizar la variable dependiente en torno a ella, y de esta manera describirla a partir de una distribución de probabilidad.

El análisis de regresión permite hacer predicciones, lo que contribuye a las técnicas de aprendizaje automático. En el caso de la inteligencia artificial es muy empleado este tipo de regresiones, y se asocia con la regresión lineal y con la estimación por mínimos cuadrados, entre otras técnicas de análisis numérico.

En algunas ocasiones el análisis de regresión puede ayudar a encontrar relaciones causales, sin embargo esto no debe confundirse, pues pudiesen haber correlaciones no causales.

El análisis de regresión puede presentar dos modelos:

### 1-Regresión Lineal Simple (RLS)

### 2-Regresión Lineal Múltiple o Polinómica (RLM)

La utilización de un modelo u otro dependerá del tipo de datos y de las características propias del estudio que se desea realizar.

### La Regresión Lineal Simple

Es un modelo sencillo, en el que se pretende encontrar la predicción de una variable continua  $Y$  a partir de otra variable continua  $X$  denominada predictora. Esto quiere decir que los datos determinados por  $Y$  son elementos experimentales, mientras que los datos de  $X$  pueden ser establecidos por el investigador bajo algún criterio, o también pueden haber sido tomados de forma experimental. En cualquiera de los casos, se considerará que la variable  $Y$  presenta aleatoriedad, mientras que la variable  $X$  carece de ella.

La formulación del modelo de RLS predice la respuesta para un valor experimental observado de  $X$  con una recta de regresión.

$$E(y|x = x) = \beta_0 + \beta_1 x \quad (25)$$

Se presenta una desviación aleatoria, denominada error aleatorio, que se denota con épsilon ( $\epsilon$ ). Debido a que se trata de un error aleatorio, se incluye en (1) y genera un modelo completo (2)



$$Y = \beta_0 + \beta_1 x + \epsilon$$

Debido a que se trata de una ecuación conocida, que es la ecuación de la recta, los parámetros a considerar para el cálculo serán la pendiente y el punto de corte de la recta con el eje vertical, que en este caso están representados por  $\beta_1$  y  $\beta_0$  respectivamente.

Sobre los errores aleatorios es posible definir lo siguiente:

Incorrelación:  $\text{Corr}(\epsilon_i, \epsilon_j) = 0$  esto significa que los valores de la respuesta o salida del sistema, no guardan relación entre si, esto significa que los valores de las respuestas no se relacionan entre ellos.

Media cero:  $E(\epsilon_i) = 0$  esto significa que hay completa dependencia de los valores de la pendiente ( $\beta_1$ ) y del punto de corte con el eje ( $\beta_0$ ) cumpliéndose una correlación lineal.

Varianza constante:  $\text{Var}(\epsilon_i) = \delta^2$  esto implica que los valores de la salida pertenecen a una misma población, donde existe una variabilidad que está determinada por  $\delta^2$ .

Para estimar el modelo se cumplen los mismos criterios clásicos para la estimación de una recta, donde las incógnitas están dadas por la pendiente y el punto de corte, y su cálculo está estrechamente relacionado con los datos obtenidos.

Existen dos métodos establecidos para determinar esta recta de regresión, que pueden ser el método de los mínimos cuadrados y el método de máximo-verosímil para cuando se desee un estudio más profundo con fines inferenciales.

### **Bondad de ajuste del modelo de Regresión Lineal**

Una vez realizado un modelo de regresión lineal, resulta necesario validar la efectividad del modelo, y garantizar que dicho modelo ofrece una predicción apropiada de la respuesta. Esta bondad de ajuste puede realizarse con las siguientes medidas:

- El error residual estimado
- El test F de bondad (obtenida a partir de la tabla ANOVA)
- El coeficiente de determinación  $R^2$

Explicaremos brevemente en qué consiste cada una de ellos,

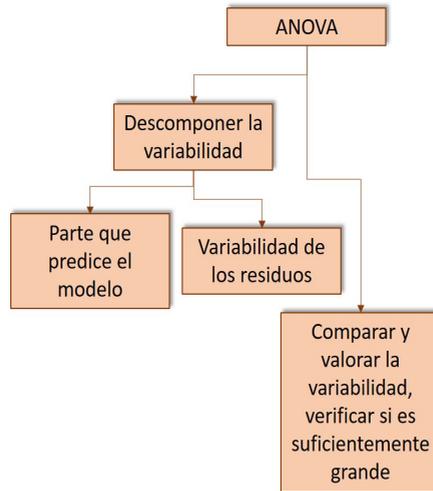
#### **1-El error residual**

Está relacionado con la escala de valores utilizado, y lo ideal es que el error sea lo menor posible, es decir, un error residual S, donde  $S^2$  viene dado por la estimación de la varianza del modelo,  $\sigma^2$ .

## Tabla ANOVA

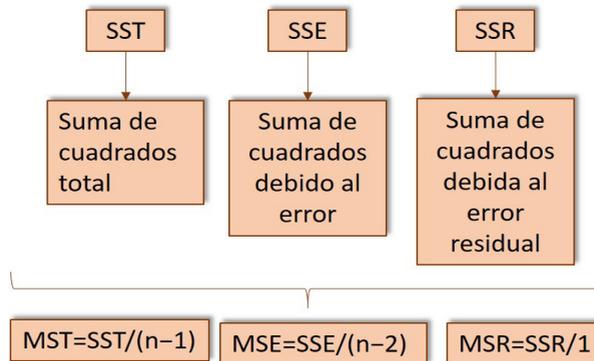
Un modelo se considera bueno si existe poca diferencia entre los datos y las predicciones, es decir, si la variabilidad logra ser explicada por el modelo.

Una tabla ANOVA es un análisis de varianza, y para construirla es necesario observar los pasos de la figura 16.



**Figura 17. Elementos básicos para la construcción de la tabla ANOVA.**

A partir de la suma de cuadrados total (SST), la suma de cuadrados debido al error residual (SSE) y la suma de cuadrados debido a la regresión (SSR) es posible determinar la variabilidad total, la predicción de la variabilidad y la variabilidad residual, tomando cada una de estas sumas y dividiéndola por su respectivo grado de libertad, como se muestra en la figura 17. Se observa que este cálculo arroja los cuadrados medios asociados.



**Figura 18. Relaciones de la suma de cuadrados y la obtención de los cuadrados medios asociados.**

Esta validación de la recta de regresión es útil para contrastar las hipótesis:

$H_0$ : indica que el modelo desarrollado no predice la respuesta de forma adecuada.

$H_1$ : indica que el modelo lineal ofrece una apropiada predicción de la respuesta.

2.El coeficiente de determinación  $R^2$

Es utilizado para determinar la bondad del ajuste de la recta de regresión, y se obtiene a partir de la descomposición de la varianza:

$$R^2 = \frac{SSR}{SST} \quad (26)$$

Este valor del coeficiente deberá estar en el rango de  $0 \leq R^2 \leq 1$ , lo que indica que cuando el valor esté próximo a 1, esto es que sea mayor de 0,6 puede considerarse que la recta de regresión realiza una buena predicción de la varianza. Por otro lado con el valor del coeficiente es cercano a cero, indicará que la variabilidad de los datos no puede ser sustentado por la recta de regresión.

### **Modelo de Regresión Lineal Múltiple o Polinómico.**

Los Modelos de Regresión Lineal Múltiples (RLM) o Modelos Polinómico (MP) presentan una diferencia con relación al modelo de RLS, y consiste en que en estos intervienen mínimo dos variables para predecir el comportamiento de la respuesta. Este modelo puede presentarse de dos tipos:

#### 1-Modelos de Regresión Lineal Múltiples

Se presentan cuando a través de una función lineal se intenta predecir una variable continua a través de un conjunto de variables continuas. Por tanto el modelo está descrito por

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon \quad (29)$$

Resulta evidente que las incógnitas de este modelo son  $\beta_0, \beta_1, \beta_2, \dots, \beta_n, \sigma^2$

Donde  $\sigma^2$  es la varianza residual del modelo.

Obsérvese que los errores se distribuyen de forma individual, y que existe una distribución normal de media cero.

$\beta_0$  describe la intercepción de la respuesta con el eje vertical, que representa la respuesta cuando la variable predictora es cero.

$\beta_i$  representan las pendientes de las rectas asociadas con variable predictora.

### Modelos de Regresión Polinómicos (RP).

Estos modelos surgen cuando en vez de intentar una predicción con una función de regresión lineal, la intentamos con una función polinómica. Este tipo de tratamiento es útil cuando se observa una tendencia curva entre las variables, entre las predictoras y la respuesta. Además permiten también realizar tratamientos con datos no lineales.

Para estos casos el modelo viene dado por:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \dots + \beta_n X^n + \epsilon \quad (28)$$

Donde:

$\beta_0$  describe la intercepción de la respuesta con el eje vertical, que representa la respuesta cuando la variable predictora es cero.

$\beta_i$  representan las pendientes de las rectas asociadas con variable predictora.

### Análisis multivariado

Se utiliza para encontrar la relación entre las variables de un conjunto de datos, con el fin de encontrar una estructura que los relacione, y a su vez conseguir la verificación de esos datos que conforman dicha estructura.

Existen distintos métodos de análisis de multivariante, entre los que es posible mencionar:

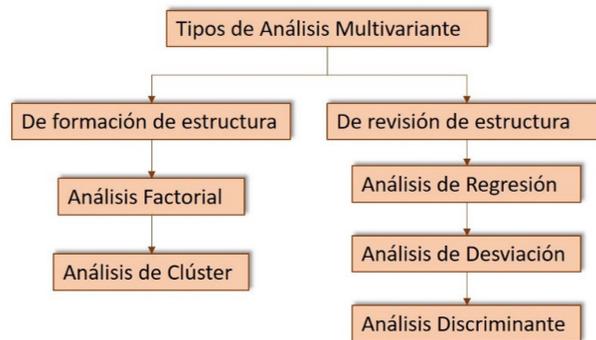


Figura 19. El análisis multivariante

**Análisis factorial:**

Consiste en la evaluación individual de los datos, por lo que está centrado en todas las variables involucradas. Se sub divide en análisis de componentes y análisis de correspondencia.

**Análisis de clúster:**

Consiste en asignar los datos a grupos de variables individuales, y clasificarlos gráficamente en función a esta asignación.

**Análisis de regresión:**

Consiste en encontrar la relación entre dos variables, una dependiente y otra no dependiente. Por un lado existe una variable predictora y por el otro existe una variable que puede predecirse.

**Análisis de desviaciones:**

Consiste en calcular los promedios de grupos de variables o de variables individuales. Por tanto es posible analizar grupos de variables, pero también es posible analizar variables individuales dentro de cada grupo.

**Análisis discriminante:**

Consiste en diferenciar grupos que pueden ser categorizados como grupos similares.

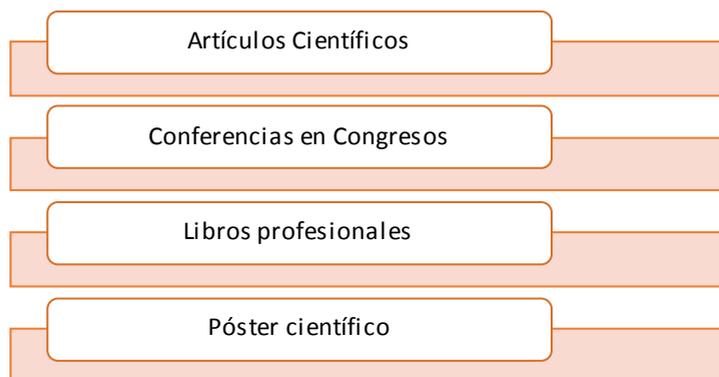
# Capítulo 7: Comunicación de la investigación cuantitativa

## Introducción

La comunicación es una parte esencial de todo proceso de investigación, pues sin ella no es posible afirmar que los objetivos fueron alcanzados o que se ejecutaron las actividades pertinentes, o simplemente no es posible afirmar o contradecir ningún aspecto producto de la investigación.

Así pues, la divulgación de los resultados de una investigación es necesaria para dar a conocer todos los aspectos que nacen producto de un proceso riguroso de trabajo científico, que abrirá posiblemente nuevos proyectos, nuevos escenarios, nuevos investigadores interesados en la temática de estudio. Al mismo tiempo que aportará a otras investigaciones.

Existen diferentes formas de divulgar un proyecto de investigación:



**Figura 19. Diferentes formas de divulgación científica**

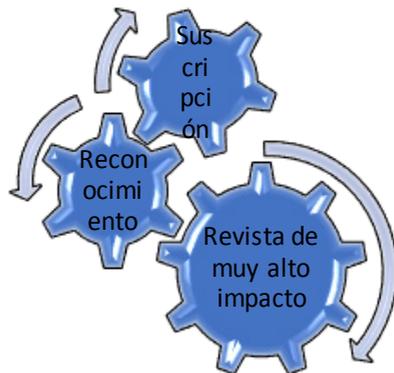
Para realizar una apropiada divulgación científica, es necesario que los autores/investigadores tengan claridad con los objetivos que se pretenden con la investigación, de manera que sea posible realizar una adecuada selección del tipo de herramienta para difundir los resultados de la investigación.

En este capítulo se darán a conocer todos los aspectos básicos necesarios que ofrecen las herramientas de divulgación científica, y que los investigadores deben tener en cuenta para poder ubicar de forma efectiva.



## Artículos Científicos

Los artículos científicos son una buena opción para divulgar los resultados de un trabajo de investigación, sin embargo no siempre resulta sencillo ubicar una revista científica apropiada para el trabajo. A la hora de seleccionar una revista científica es necesario que los autores evalúen las posibilidades que ofrece su trabajo, si los autores son expertos en redacción y si la investigación tienen un alto impacto científico, podría ser oportuno para ellos optar a publicaciones en revistas de muy alto nivel, con un alto factor de impacto. Sin embargo esto puede traer ciertos inconvenientes, en la figura 20 es posible apreciar algunas desventajas y ventajas en la selección de revistas científicas.



**Figura 20. Ventajas y desventajas de las revistas de alto impacto**

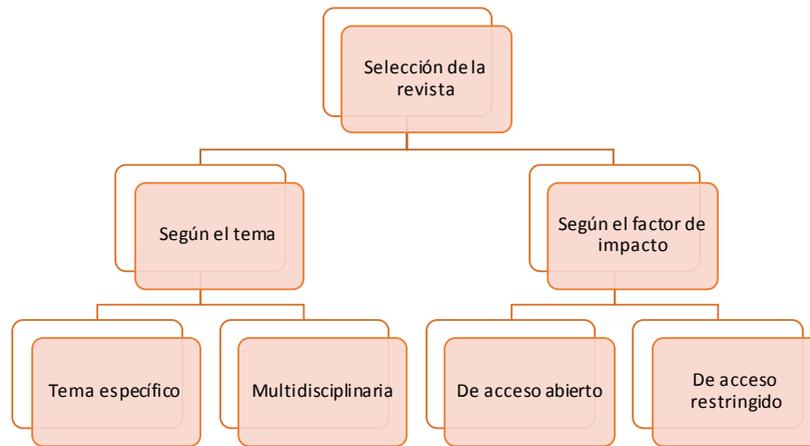
Las revistas de alto impacto pueden ofrecer un excelente prestigio para el artículo científico, esto debido a que suelen tener un importante comité técnico entrenado y con muy alto prestigio profesional, que evalúa las publicaciones, corrige y permite ofrecer una alta calidad del artículo. Sin embargo, su principal desventaja es que la gran mayoría de las revistas de alto impacto no tienen acceso abierto, sino que solo aquellas personas con suscripción pueden acceder a las publicaciones, lo que restringe bastante la visibilidad del artículo.

Normalmente se suele pensar que las revistas de alto impacto pueden darle más realce a nuestro trabajo, pero no necesariamente debe ser así, ya que al tener restricciones de acceso, muchas veces los mismos investigadores no pueden acceder a la publicación de su propio trabajo.

Por otro lado se encuentran las revistas de acceso abierto, este tipo de revistas no necesariamente es de alto impacto, muchas veces tienen menor impacto, pero poseen altísima calidad en el proceso de selección de los trabajos. Lo que hace que el trabajo esté disponible al público y sea más accesible la información para otros centros de investigaciones.

Las revistas de acceso abierto nos permiten realizar publicaciones de un trabajo de investigación, y también permiten acceder a muchos trabajos de investigación que pueden enriquecer los trabajos desarrollados. Hoy en día existen redes sociales de investigadores, que también pueden servir para compartir ciertos aspectos de una investigación, y al mismo tiempo contribuir con otros proyectos.

La selección de una revista científica también debe hacerse considerando la temática del estudio, pues esto permitirá incluir nuestro trabajo en un público selectivo con altos niveles profesionales en el tema, que pueden enriquecerse con el proyecto y que pueden interesarse en el tema y crear redes de investigadores o trabajos en conjunto.



**Figura 21. Aspectos a considerar para la selección de una revista**

Las revistas multidisciplinarias ofrecen una excelente opción para publicar un trabajo, pero es necesario que se tome en cuenta la calidad de evaluación de un trabajo, para que pueda ser un trabajo bien visto y pueda ser seleccionado por los autores relacionados a la temática.

### **Conferencias en congresos**

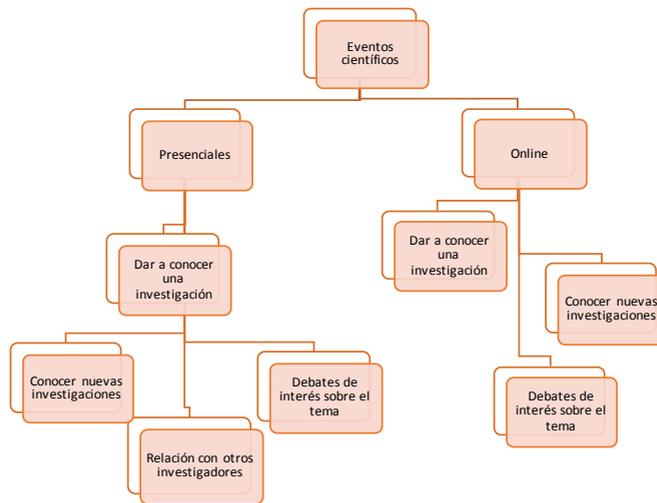
Otra de las formas de hacer conocer un trabajo de investigación, es a través de eventos científicos y académicos. Este tipo de eventos permite un trato directo con la comunidad científica, y permite darnos a conocer como investigadores.

Los congresos y eventos científicos, son una oportunidad para generar debates académicos, encontrar mejoras para el trabajo, encontrar falencias y destacar las fortalezas. Así como también es oportuno para relacionarnos con otros colegas de la misma área, que puedan contribuir o darnos ideas para futuros proyectos.

Los congresos online son una buena opción para dar a conocer un trabajo, más aún si estos tienen incorporados publicaciones de memorias o artículos científicos. En todos los casos, los eventos académicos son buenas opciones para divulgar un trabajo, ayudan de muchas maneras para la visualización del trabajo y para la valoración de los investigadores.

Un buen trabajo de investigación siempre es reconocible a simple vista por los investigadores del área, por tanto, es necesario que nuestro trabajo tenga un buen impacto visual y científico. La relación en espacios de debate con otros investigadores también debe ser la más apropiada, ya que eso puede abrir oportunidades y nuevos proyectos de investigación. La sencillez y la humildad suelen ser los valores más significativos de un buen investigador.

Un espacio de debate científico en torno a nuestro tema de investigación, por lo general produce buenos caminos para futuros proyectos, además permite ver las falencias y fortalezas de nuestro trabajo. También es posible incluir las relaciones inter institucionales, inter disciplinarias y demás que resulten para el beneficio de nuestra trayectoria como investigadores.



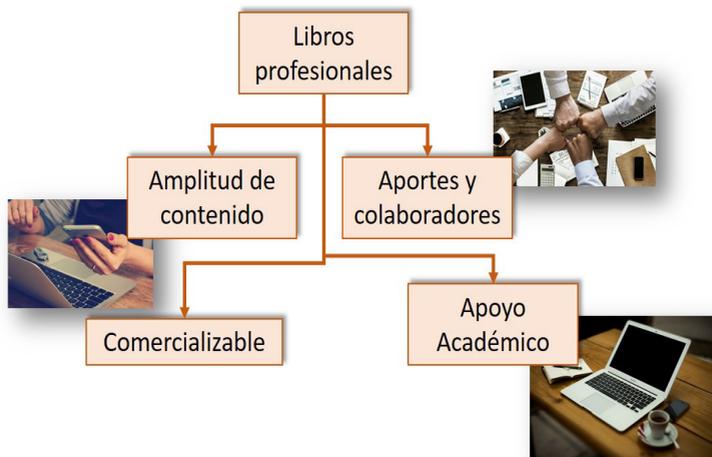
**Figura 22. Un evento científico siempre es beneficioso para la visualización de nuestro trabajo, sea online o presencial.**

### Libros profesionales

Otra forma de hacer visible en la comunidad científica un trabajo de investigación, es a través de libros profesionales, ya que pueden elaborarse a partir de los resultados de una investigación, y pueden facilitar la comprensión de las ideas de nuestro trabajo, ya que el contenido es mucho más extenso que un artículo científico.

Los libros pueden incluir capítulos relevantes de un trabajo de investigación, pueden ofrecerse en centros de estudios y librerías, y permiten que los autores destaquen en la comunidad científica. Además es posible hacer segundas ediciones, mejores aportes, trabajarlos digital o impreso.

Todas las publicaciones científicas a través de libros pueden ser fructíferas para la comunidad de investigadores, y pueden aportar grandemente a las universidades donde pertenecen, aportando a las materias específicas y promoviendo el desarrollo local de la obra.



**Figura 23. Los libros profesionales ofrecen importantes aportes al grupo de investigadores.**

Los libros profesionales deben publicarse con editoriales especializadas, pues la calidad de un libro puede asegurar la trayectoria del mismo, pero una mala calidad puede arruinar la imagen de los investigadores y disminuir la posibilidad de uso del libro, su venta y su aplicación en ambientes de enseñanza.

Una buena editorial, tomará en cuenta los detalles ortotipográficos, pero además los aspectos estéticos, el software de edición especializado, el diseño apropiado de carátulas, y sin duda la revisión por pares especializados.

Por otro lado una editorial poco cuidadosa, dejará pasar la calidad de la obra y podría arruinar la buena reputación de los autores. También es importante que el contenido del libro sea cautivador para un público objetivo, con alta calidad científica y con carácter académico y profesional, que permita divulgar los resultados de investigación de forma atractiva y profesional.

### **Póster científico**

Los pósteres científicos son herramientas de divulgación científica, que contienen los aspectos más relevantes de una investigación. No debe confundirse con un cartel publicitario, pues la idea no es comercializar un producto sino dar a conocer los resultados de una investigación, con carácter académico.

El póster científico debe contener la esencia de una investigación y debe poder transmitir una idea y un conjunto de resultados. Debe destacar los nombres de los autores, y deben resaltar las imágenes asociadas



al trabajo de investigación.

El desarrollo del texto de un póster científico debe ser resumido y con una tipografía acorde al diseño, que sea de fácil lectura y que sea armonioso con el diseño del póster. El diseño dependerá de las exigencias de los organizadores del evento académico.

Los elementos de un póster pueden variar según el tipo de evento, según el tipo de trabajo, pero en la figura 24 podemos ofrecer algunas ideas generales de los elementos más esenciales de un póster.

Al igual que en un artículo científico el póster científico, debe contener la identificación del póster según las características del evento, además debe contener el título del trabajo de investigación, para que pueda distinguirse entre los demás y capturar la atención de los especialistas del área.

Los aspectos básicos de un póster incluyen los mismos aspectos que otras publicaciones científicas, partiendo de la introducción, objetivos, métodos, resultados y conclusiones, pero estos deben hacerse de forma sintetizada y simplificada, que puedan capturar la atención del público sin saturar al lector con largos escritos, siempre apoyándose en figuras, diagramas, y esquemas que enriquezcan la información.

La visualización de los autores debe ser apropiada para destacar el aporte científico sin presumirle al público, por ende debe hacerse de manera elegante y sencilla. Las figuras y el texto no pueden ser grotescas ni causar incomodidad visual a los presentes. Debe ser armoniosa con el tema de estudio.

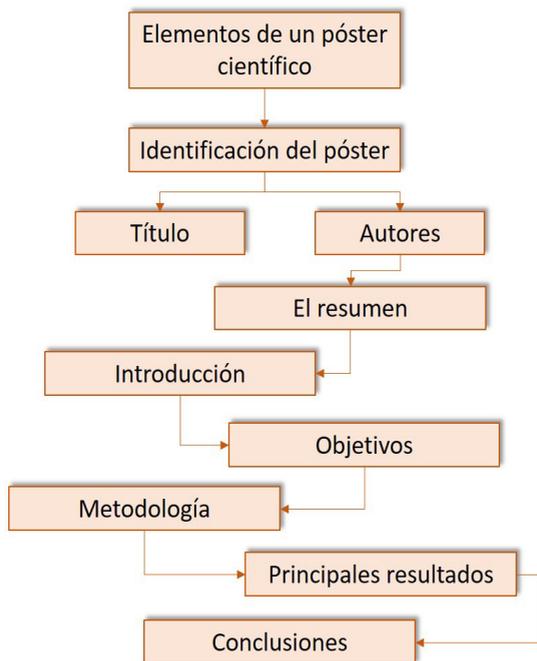


Figura 24. Los elementos del poster pueden variar, pero los aspectos más básicos se destacan en esta figura, resaltando que el trabajo de investigación puede exigir otras características en el póster.

## REFERENCIAS

- Briones, G., (2002). Metodología de la investigación cuantitativa en las ciencias sociales, Bogotá Colombia: INSTITUTO COLOMBIANO PARA EL FOMENTO DE LA EDUCACIÓN SUPERIOR, ICFES.
- McGraw-Hill (Ed.). (2014). Metodología de la investigación. México D. F., México: McGRAW-HILL/INTERAMERICANA EDITORES, S.A. DE C.V.
- Monje, C., (2011). Metodología de la investigación cuantitativa y cualitativa una guía didáctica, Colombia, Neiva: UNIVERSIDAD SURCOLOMBIANA FACULTAD DE CIENCIAS SOCIALES Y HUMANAS PROGRAMA DE COMUNICACIÓN SOCIAL Y PERIODISMO.
- Universidad Pedagógica Experimental Libertador. (2016). Manual de Trabajos de Grado de Especialización y Maestría y Tesis Doctorales. Caracas, Venezuela: Fondo Editorial de la Universidad Pedagógica Experimental Libertador.
- López-Roldán, P., Fachelli, S. (2015). Metodología de la investigación social cuantitativa, Barcelona, España: Dipòsit Digital de Documents, Universitat Autònoma de Barcelona.
- Tamayo y Tamayo, M. (1999). Modulo 2. La Investigación, Serie Aprender a investigar. Santa Fé de Bogotá, Colombia: INSTITUTO COLOMBIANO PARA EL FOMENTO DE LA EDUCACIÓN SUPERIOR.
- Posada, G. (2016). Elementos básicos de estadística descriptiva para el análisis de datos, Medellín, Colombia: Fundación Universitaria Luis Amigó.
- Campbell, D. T., Stanley, J. C. (1995). Diseños experimentales y cuasi experimentales en la investigación social, Buenos Aires, Argentina: Amorrortu Editores.
- Canales C., M. (Ed). (2006). Metodologías de investigación social, introducción a los oficios, Santiago de Chile, Chile: LOM Editores.
- Ramos, C. (2017). Los paradigmas de la investigación científica. Avances En Psicología, 23(1), 9-17. <https://doi.org/10.33539>.
- Hurtado, J. (2005). Cómo formular objetivos de investigación, Caracas, Venezuela: Quirón Ediciones Fundación Sypal. Instituto Universitario de Tecnología Jose Antonio Anzoategui.
- Colectivo de autores (2002). Métodos y técnicas aplicados a la investigación en atención primaria de salud. La Lisa, Cuba: Finlay/Ediciones.
- Blaxter, L., Hughes, C., y Tight, M. (2000). Cómo se hace una investigación. Barcelona: Gedisa.
- Salinas, P., Cárdenas, M. (2009). Métodos de investigación social. Quito, Ecuador: Editorial Quipus, CIESPAL.



Sarmentero, I., Falcón, O. y Portuondo, A. Metodología para la determinación de indicadores cuantitativos y su evaluación en la asimilación de los valores organizacionales. *Avanzada Científica* (11)2. Pp 1-9.

# Validez y Confiabilidad de Instrumentos de Investigación



## Validez y Confiabilidad de Instrumentos de Investigación

### Validez

La validez responde a la pregunta ¿con qué fidelidad corresponde la población al atributo que se va a medir?, la validez de un instrumento consiste en que mida lo que tiene que medir (autenticidad), algunos procedimientos a emplear son: Know groups (preguntar a grupos conocidos), predictive validity (comprobar comportamiento) y cross-checkquestions (contrastar datos previos); al estimar la validez es necesario saber a ciencia cierta qué rasgos o características se desean estudiar, a este rasgo o característica se le denomina variable criterio: Existen tres tipos de validez:

Validez de Contenido: se refiere al grado en que un instrumento refleja un dominio específico del contenido de lo que se quiere medir, se trata de determinar hasta dónde los ítems o reactivos de un instrumento son representativos del universo de contenido de la característica o rasgo que se quiere medir, responde a la pregunta cuán representativo es el comportamiento elegido como muestra del universo que intenta representar. por ejemplo, un cuestionario sobre la actitud de los alumnos ante la investigación no tendrá validez de contenido si explora la opinión de los alumnos sobre las características de los docentes dentro de la cátedra de estadística, también se le denomina validez racional o lógica.

El análisis del instrumento se hace en gran parte en términos de su contenido. Sin embargo, no se debe pensar en el contenido de manera estrecha, porque puede ser que estemos interesados en un proceso tanto como en el contenido simplemente, el problema de apreciar la validez de contenido está vinculado íntimamente con la planificación del cuestionario y después con la construcción de los ítems ajustados a esos planes y a los contenidos del marco teórico de la investigación, un instrumento de medición debe tener representados a todos los ítems del dominio de contenido de las variables a medir.

Los investigadores deben elaborar una serie de ítems, acordes con las variables empleadas y sus respectivas dimensiones, luego de la selección de los ítems más adecuados para el proyecto, se elabora el instrumento, para ser validado por un grupo impar de expertos, normalmente de tres o cinco, que certifiquen, efectivamente, que las preguntas, reactivos o afirmaciones seleccionadas son claras y tienen coherencia con el trabajo desarrollado.

Hay que considerar que, la validez de contenido no puede expresarse cuantitativamente es más bien una cuestión de juicio, se estima de manera subjetiva o intersubjetiva empleando, usualmente, el denominado Juicio de Expertos. Se recurre a ella para conocer la probabilidad de error probable en la configuración del instrumento, mediante el juicio de expertos se pretende tener estimaciones razonablemente buenas, las «mejores conjeturas». Sin embargo, estas estimaciones pueden y deben ser confirmadas o modificadas a lo largo del tiempo, según se vaya recopilando información durante el funcionamiento del sistema, los juicios de expertos se pueden obtener por métodos grupales o por métodos de experto único. Se pueden seguir, entre otros, el método de agregados individuales, el método Delphi, la técnica de grupo nominal, v de Aiken y el método de consenso grupal.

Se procede de la siguiente manera:

Se seleccionan al menos tres expertos o jueces, para juzgar de manera independiente la relevancia y congruencia de los reactivos con el contenido teórico, la claridad en la redacción y el sesgo o tendenciosidad en la formulación de los ítems, es decir, si sugieren o no una respuesta.

Cada experto debe recibir la información escrita suficiente sobre: el propósito de la prueba (objetivos), conceptualización del universo de contenido, tabla de especificaciones o de operacionalización de las variables del estudio.

Cada experto debe recibir un instrumento de validación que contenga: congruencia ítem-dominio, claridad, tendenciosidad o sesgo y observaciones.

Se recogen y analizan los instrumentos de validación y se decide:

1) los ítems que tienen 100% de coincidencia favorable entre los jueces (congruentes, claros en su redacción y no tendenciosos) quedan incluidos en el instrumento

2) los ítems que tengan 100% de coincidencia desfavorable entre los jueces quedan excluidos del instrumento

3) los ítems que tengan una coincidencia parcial entre los jueces deben ser revisados, reformulados o sustituidos, si es necesario, y nuevamente validados.

**Cuadro 1. Formato para validar instrumentos a incluir en el instrumento de validación.**

Ítem	CRITERIOS A EVALUAR										Observaciones (si debe eliminarse o modificarse un ítem por favor indique)	
	Claridad en la redacción		Coherencia interna		Inducción a la respuesta (Sesgo)		Lenguaje adecuado con el nivel del informante		Mide lo que pretende			
	Sí	No	Sí	No	Sí	No	Sí	No	Sí	No		
1												
2												
3												
...												
n												



Aspectos Generales		Sí	No	*****
el instrumento contiene instrucciones claras y precisas para responder el cuestionario				
los ítems permiten el logro del objetivo de la investigación				
los ítems están distribuidos en forma lógica y secuencial				
el número de ítems es suficiente para recoger la información. en caso de ser negativa su respuesta, sugiera los ítems a añadir				
<b>Validez</b>				
aplicable			no aplicable	
aplicable atendiendo a las observaciones				
Validado por:		c.i.:		Fecha:
Firma:		teléfono:		e-mail:

**Validez de Constructo:** intenta responder la pregunta ¿hasta dónde el instrumento mide realmente un rasgo determinado y con cuánta eficiencia lo hace?, al respecto esta validez interesa cuando se quiere usar el desempeño de los sujetos con el instrumento para inferir la posesión de ciertos rasgos, para estudiar este tipo de validez es necesario que exista una conceptualización clara del rasgo estudiado basado en una teoría determinada, la teoría sugiere las tareas pruebas que son apropiadas para observar el atributo o rasgo y las evidencias a considerarse en la evaluación, se sugiere los siguientes pasos:

1) Identificar las construcciones que pudieran explicar la ejecución en el instrumento

2) Formulación de hipótesis comprobables a partir de la teoría

3) Recopilación de los datos para probar las hipótesis el término constructo se usa en psicología para referirse a algo que no es observable, pero que literalmente es construido por el investigador para resumir o explicar las regularidades o relaciones que él observa en la conducta. por tanto, la mayoría de los nombres de rasgos se refieren a constructos, para las preguntas acerca de si el instrumento revela algo significativo respecto de las personas, se usa el término validez de constructo.

Muchas pruebas psicológicas, y en menor medida algunas pruebas educativas, intentan medir rasgos generales o cualidades de un individuo, tales como: razonamiento verbal, visualización espacial, sociabilidad, introversión e interés mecánico son designaciones de constructos o de rasgos, las pruebas de estas funciones son válidas en tanto se comporten de la manera que cabría esperar que se comportara el rasgo.



**Validez Predictiva o de Criterio Externo o Empírica:** se asocia con la visión de futuro, determinar hasta dónde se puede anticipar el desempeño futuro de una persona en una actividad determinada, la validez predictiva se estudia comparando los puntajes de un instrumento con una o más variables externas denominadas variables de criterio. Se establece una correlación, la cual se interpreta como índice de validez.

Para este propósito, estamos interesados en el grado en que el instrumento se correlaciona con algunos de los criterios escogidos para medir, por ejemplo: el éxito en los estudios, así, se toma alguna otra medida, como el criterio del “éxito”, y juzgamos el cuestionario elaborado en términos de su relación con esa medida de criterio. El procedimiento básico consiste en aplicar la prueba a un grupo de personas que ingresan a un trabajo o a un programa de entrenamiento y, posteriormente, seguir las observando para obtener de cada una, una medida criterio de éxito específica y luego calcular la correlación entre la puntuación de prueba y la medida criterio de éxito. Cuanta más alta sea la correlación, mejor será el instrumento, es decir, la evaluación como predictor es principalmente una evaluación empírica y estadística.

**Cualidades deseables en una medida criterio:** Existen cuatro cualidades deseables en una medida criterio, en orden de importancia son:

-Atingencia – Se consideran que un criterio es atingente en la medida en que esta medida criterio corresponde con o ejemplifica el éxito en el trabajo.

-Libre de sesgos – la medida debería ser aquella en la que cada persona tiene las mismas oportunidades de obtener una buena puntuación.

-Confiable – debe ser estable o reproducible si ha de ser predicha por algún tipo de prueba.

-Disponibilidad – debe tener en cuenta límites prácticos como: ¿Qué tanto se debe esperar para obtener una puntuación?, ¿cuánto dinero costarán las interrupciones de las actividades?

Factores que Afectan la Validez

Existen varios factores que tienden a distorsionar los coeficientes de validez y a complicar su interpretación, entre ellos:

Construcción del Instrumento. algunos de los factores a evitar, para que el instrumento sea válido son:

-Instrucciones imprecisas o vagas

-Estructura sintáctica de las oraciones muy difíciles

-Preguntas inadecuadas respecto a las especificaciones (sobre todo en pruebas escritas y orales)

-Preguntas que sugieren la respuesta

-Ambigüedad en la formulación de los reactivos, que lleven a diferentes interpretaciones



-Cuestionarios demasiado cortos (no incluyen una muestra adecuada de los indicadores a medir), con pocos reactivos

-Ítems incongruentes con el universo de contenido, sin relación con los rasgos o características a medir

-Ordenamiento inadecuado de los ítems

-Patrón identificable de respuestas (sobre todo en pruebas de selección)

**Administración y Calificación del Instrumento.** Entre estos factores están: tiempo insuficiente para responder, ayuda adicional de otros sujetos, en caso de pruebas: las chuletas o copias, la subjetividad en los puntajes de las preguntas.

**Respuestas de los Sujetos.** Algunos factores: bloqueo de los sujetos al responder debido a situaciones emocionales y las respuestas al azar.

**Naturaleza del Grupo y del Criterio.** Debe ser aplicado el instrumento para el grupo al cual fue diseñado, puede ser inapropiado para otros grupos.

## Confiabilidad

Antes de iniciar el trabajo de campo, es imprescindible probar el cuestionario sobre un pequeño grupo de población. Esta prueba piloto ha de garantizar las mismas condiciones de realización que el trabajo de campo real. Se recomienda un pequeño grupo de sujetos que no pertenezcan a la muestra seleccionada pero sí a la población o un grupo con características similares a la de la muestra del estudio, aproximadamente entre 14 y 30 personas, de esta manera se estimará la confiabilidad del cuestionario.

La confiabilidad responde a la pregunta ¿con cuánta exactitud los ítems, reactivos o tareas representan al universo de donde fueron seleccionados?, el término confiabilidad "...designa la exactitud con que un conjunto de puntajes de pruebas mide lo que tendrían que medir". Entre los métodos para estimar la confiabilidad, se tienen:

Método Test-Retest: una forma de estimar la confiabilidad de un test o cuestionario es administrarlo dos veces al mismo grupo y correlacionar las puntuaciones obtenidas. Este método tiene la desventaja de que los puntajes pueden verse afectados por el recuerdo, la práctica, etc. este procedimiento no es adecuado para aplicarlo a pruebas de conocimientos sino para la medición de aptitudes físicas y atléticas, tests de personalidad y motores. El coeficiente que se obtiene recibe el nombre de coeficiente de estabilidad porque denota la coherencia de las puntuaciones en el tiempo, para un desarrollo adecuado y sean confiables deben variar entre 0,80 y 0,95.

Método común de división por mitades o Hemitest: este método computa el coeficiente de correlación entre los puntajes de las dos mitades del test o cuestionario aplicado, esto supone que los dos test mitades son paralelos, tienen igual longitud y varianza entre sí. Se estima a través del coeficiente de confiabilidad de Spearmanbrown:

Se establece la correlación entre los dos puntajes de las dos mitades del test a través del método de los puntajes directos, Correlación  $r$  de Pearson:

**Coefficiente Alfa de Cronbach:** para evaluar la confiabilidad o la homogeneidad de las preguntas o ítems es común emplear el coeficiente alfa de Cronbach cuando se trata de alternativas de respuestas policotómicas, como las escalas tipo likert; la cual puede tomar valores entre 0 y 1, donde: 0 significa confiabilidad nula y 1 representa confiabilidad total, el coeficiente  $\alpha$  de Cronbach puede ser calculado por medio de la consistencia interna del instrumento.

El coeficiente omega, a diferencia del coeficiente de alfa trabaja con las cargas factoriales, que son la suma ponderada de las variables estandarizadas, transformación que hace más estable los cálculos y refleja el verdadero nivel de fiabilidad. En segundo lugar, no depende del número de ítems tal como se aprecia en su expresión matemática. En tercer lugar, el coeficiente omega es considerado una adecuada medida de la confiabilidad si no se cumple el principio de tal equivalencia, el cual puede incumplirse si los coeficientes de los ítems que conforman una matriz de solución factorial presentan valores muy diferentes. Para considerar un valor aceptable de confiabilidad mediante el coeficiente omega, éstos deben encontrarse entre .70 y .90, aunque en algunas circunstancias pueden aceptarse valores superiores a .65.

Interpretación del Coeficiente de Confiabilidad. El coeficiente de confiabilidad es un coeficiente de correlación, teóricamente significa la correlación del test consigo mismo. Sus valores oscilan entre 0 y 1. entre las escalas empleadas se tienen las mostradas en los cuadros siguientes.

**Cuadro 2. Escala de interpretación de la magnitud del coeficiente de correlación o asociación.**

coeficiente de correlación	magnitud
0,70 a 1,00	muy fuerte
0,50 a 0,69	Sustancial
0,30 a 0,49	moderada
0,10 a 0,29	baja
0,01 a 0,09	despreciable

**Cuadro 3. Interpretación de la magnitud del coeficiente de confiabilidad de un instrumento.**

rangos	magnitud
0,81 a 1,00	muy alta
0,61 a 0,80	alta
0,41 a 0,60	moderada
0,21 a 0,40	baja
0,01 a 0,20	muy baja



**Cuadro 4. Interpretación de la magnitud del coeficiente de confiabilidad de un instrumento.**

<b>Correlación</b>	<b>Directa</b>	<b>Indirecta</b>
alta	0,70 a 1,00	-1,00 a -0,70
moderada	0,40 a 0,70	-0,70 a -0,40
baja	0,10 a 0,40	-0,10 a -0,10
nula	-0,10 a 0,10	-0,10 a 0,10

Finalmente, es importante señalar que las normas de interpretación presentadas en los cuadros se encuentran en los promedios estandarizados, siendo el caso para asumir que un instrumento es confiable la medición debe estar por encima de 0,7.

#### **Algunos ejemplos de instrumentos validados por los investigadores**

<b>Escala de salud mental positiva</b>		F1	F2	F3
17	Intento mejorar como persona	0.66		
22	Soy capaz de mantener un buen nivel de autocontrol en las situaciones conflictivas de mi vida	0.59		
27	Cuando hay cambios en mi entorno intento adaptarme	0.58		
21	Soy capaz de controlarme cuando tengo pensamientos negativos	0.56		
16	Intento sacar los aspectos positivos de las cosas malas que me suceden	0.56		
20	Creo que soy una persona sociable	0.54		
28	Delante de un problema soy capaz de solicitar información	0.53		
15	Soy capaz de tomar decisiones por mí misma/o	0.52		
23	Pienso que soy una persona digna de confianza	0.52		
26	Si estoy viviendo presiones exteriores desfavorables soy capaz de continuar manteniendo mi equilibrio personal	0.5		
25	Pienso en las necesidades de los demás	0.46		
4	Me gusto como soy	0.44		
36	Cuando se me plantea un problema intento buscar posibles soluciones	0.44		
18	Me considero "un/a buen/a psicólogo/a", las personas vienen a cuando tienen problemas	0.43		
11	Creo que tengo mucha capacidad para ponerme en el lugar de los demás y comprender sus respuestas	0.42		
32	Trato de desarrollar y potenciar mis buenas aptitudes	0.4		
14	Me considero una persona menos importante que el resto de las personas que me rodean		0.8	
31	Creo que soy un/a inútil y no sirvo para nada		0.78	
38	Me siento insatisfecha/o conmigo misma/o		0.74	

12	Veó mi futuro con pesimismo	0.71
19	Me preocupa que la gente me critique	0.68
13	Las opiniones de los demás me influyen mucho a la hora de tomar mis decisiones	0.61
34	Cuando tengo que tomar decisiones importantes me siento muy insegura/o	0.56
10	Me preocupa mucho lo que los demás piensan de mí	0.52
33	Me resulta difícil tener opiniones personales	0.47
7	Para mí, la vida es aburrida y monótona	0.46
1	Me resulta especialmente difícil aceptar a los otros cuando tienen actitudes distintas a las mías	0.84
8	A mí, me resulta especialmente difícil dar apoyo emocional	0.83
2	Los problemas me bloquean fácilmente	0.79
9	Tengo dificultades para establecer relaciones interpersonales profundas y satisfactorias con algunas personas	0.75
3	Me resulta especialmente difícil escuchar a las personas que me cuentan sus problemas	0.75
6	Me siento a punto de explotar	0.69

En la revisión del contenido de los ítems del Factor 1, se ha identificado que la mayoría de los ítems están relacionados con tres características capacidad de adaptación, valoración personal y empatía, por lo que decide llamarla capacidad de adaptación. En el Factor 2, se han incluido ítems relacionados con los factores de satisfacción personal y autonomía de la estructura original, por lo que decide llamarla satisfacción personal y autonomía. En el Factor 3, contiene información más variada, orientada con la tolerancia hacia los demás, las habilidades para relacionarse y control emocional, que se le denomina tolerancia a la frustración.

**Enlace de la publicación del artículo: Propiedades psicométricas de la escala de salud mental positiva en Arequipa (Perú)**

<https://revistas.um.es/analesps/article/view/472061>

**Revista Anales de Psicología**

**Base Scopus España**



## Escala de Clima organizacional

<b>CULTURA ORGANIZACIONAL</b>		$\lambda$
<b>Motivación</b>		
1	Mi centro de labores me ofrece la oportunidad de hacer lo que mejor sé hacer.	0.60
2	Mi jefe inmediato se preocupa por crear un ambiente laboral agradable.	0.81
3	Me brindan reconocimiento por un buen desempeño laboral.	0.54
4	Reconocen mi esfuerzo si trabajo más de las horas reglamentarias.	0.50
<b>Identidad</b>		
6	Me interesa el desarrollo de la Universidad.	0.75
7	Me siento contento de ser parte del desarrollo de la Universidad.	0.91
8	Estoy satisfecho y orgulloso de trabajar en la Universidad.	0.73
<b>Apoyo</b>		
15	En la universidad los jefes cuidan a su personal.	0.74
16	Mi jefe me respalda y deja que aprenda de mis errores.	0.77
17	Mi jefe me trata con respeto.	0.87
<b>Equidad</b>		
18	Puedo contar con un trato justo por parte de mi jefe.	0.41
19	Los directivos tratan igual a todos los que trabajan en la Universidad.	0.51
20	Mi jefe no tiene favoritismos.	0.70
21	Si mi jefe sanciona a alguien es porque probablemente esa persona se lo merece.	0.81
<b>Trabajo en equipo</b>		
26	Los objetivos están claramente establecidos en mi equipo de trabajo.	0.82
27	El apoyo y escucha son frecuentes en mi equipo de trabajo.	0.87
28	Se hace presente el intercambio de ideas en mi equipo de trabajo.	0.84
<b>Organización</b>		
30	Las labores que se desarrollan en mi área de trabajo se encuentran organizadas.	0.72
31	La documentación (procedimientos, instructivos y formatos) son útiles y apropiadas para realizar mi trabajo.	0.69
32	Conozco como se encuentra estructurada mi área de trabajo y cuáles son las actividades que se desarrollan.	0.52
<b>Condiciones de trabajo</b>		
34	En mi centro de trabajo existen accesos adecuados para el personal discapacitado.	0.70
35	Las áreas de trabajo se conservan limpias y en orden, permitiendo el desarrollo de mis actividades.	0.56
36	Dispongo de suficientes instalaciones sanitarias y áreas de descanso.	0.80
37	En general, las condiciones de trabajo de mi área son seguras.	0.61
<b>DISEÑO ORGANIZACIONAL</b>		
<b>Comunicación Organizacional</b>		$\lambda$
39	Mi jefe inmediato trata de obtener información antes de tomar una decisión.	0.54
45	Mi jefe inmediato se comunica regularmente con los trabajadores para recabar apreciaciones técnicas o percepciones relacionadas al trabajo.	0.76
46	Mi jefe inmediato me comunica si estoy realizando mi trabajo, bien o mal.	0.52
47	Presto atención a los comunicados que emiten mis jefes.	0.78
<b>Reconocimiento</b>		
41	Mi salario y beneficios son razonables.	0.31
42	Mi remuneración es adecuada en relación con el trabajo que realizo.	0.54
49	Se ofrecen posibilidades para el desarrollo personal de los trabajadores.	0.53
50	Mi jefe se fija más en las habilidades de los trabajadores que en sus debilidades.	0.54
<b>Espacio físico</b>		
52	La iluminación en mi espacio de trabajo es:	0.61
53	El servicio de aseo en mi espacio de trabajo es:	0.72
54	El equipo para realizar mi trabajo es:	0.53
56	El ruido para el desarrollo de mis actividades es:	0.42

<b>POTENCIAL HUMANO</b>		$\lambda$
<b>Innovación</b>		
57	La innovación es característica de la Universidad.	0.62
58	Mis compañeros de trabajo toman iniciativas para la solución de problemas.	0.56
59	En la Universidad todos están abiertos al cambio.	0.71
60	En la Universidad continuamente se adoptan nuevos y mejores métodos para hacer el trabajo.	0.71
<b>Liderazgo</b>		
61	Existe una clara indicación de las funciones que cada uno debe desempeñar.	0.84
62	Mi jefe inmediato me permite participar en la planificación del trabajo.	0.82
64	A menudo se inician trabajos que no se sabe por qué se hacen.	0.30
<b>Autonomía</b>		
68	Tengo autonomía laboral.	0.59
69	Tengo libertad en la toma de decisiones.	0.41
70	Organizo mi trabajo como mejor me parece.	0.49
<b>Cohesión</b>		
72	Promuevo un ambiente amigable dentro de la Universidad.	0.51
73	Existe amistad y cariño entre los miembros de la Universidad.	0.77
74	Tengo libertad para comunicarme abiertamente con mis superiores.	0.81
75	Existe buena relación entre los superiores y subordinados.	0.78

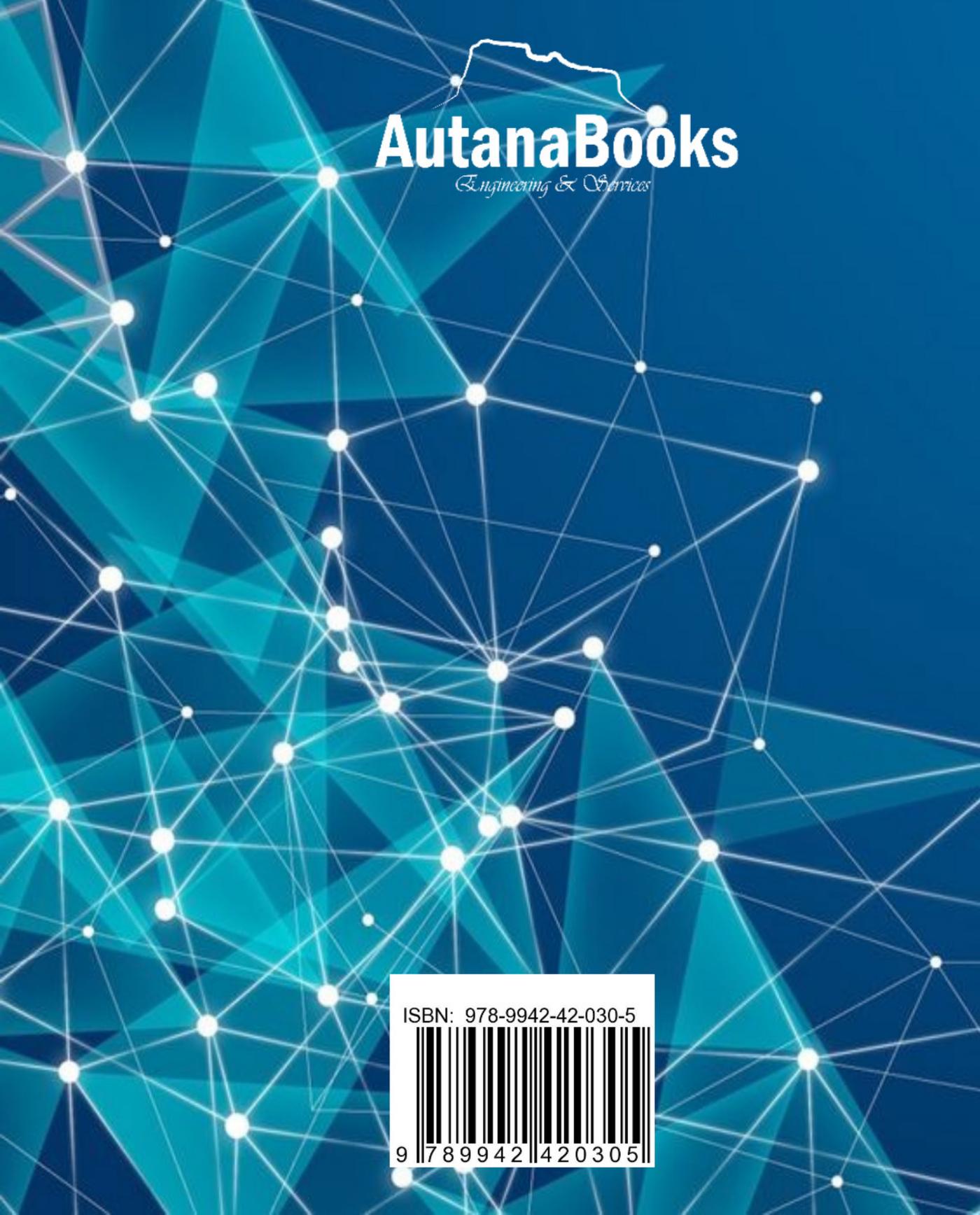
Enlace de la publicación del artículo: Análisis psicométrico de la escala clima organizacional en servidores administrativos de una universidad pública

<https://journal.universidadean.edu.co/index.php/Revista/article/view/3051>

Revista Escuela de Administración de Negocios

Base Scielo Colombia





# AutanaBooks

*Engineering & Services*

ISBN: 978-9942-42-030-5



9 789942 420305