

# UN ENFOQUE SOBRE LA INVESTIGACIÓN CIENTÍFICA EN ECONOMÍA

Douglas C. Ramírez Vera<sup>1</sup>

*“Examinadlo todo y quedaos con lo que vale” (1 Ts 5,21)*

## 1. INTRODUCCIÓN

Nadie quiere equivocarse de camino, todos quieren estar seguros de ir por el camino correcto, siempre se quiere tener la certeza de que el camino elegido, el cual se sigue sea el verdadero. Todos conocen el efecto que produce en cada uno, la sola posibilidad de la equivocación, del no saber donde se esta y a donde se va, esta certeza es necesaria para la seguridad mental, emocional y espiritual de cada hombre y mujer.

No es cuestión de cuan coherente o cuanto esfuerzo se ponga, no es el resultado del “voluntarismo” no es suficiente decir estoy seguro, es necesario estar seguro y esto es producto del juicio cierto. Pero es necesario primero aprender a juzgar. Esta certeza en el juicio permite avanzar lenta, trabajosamente, incluso cojeando pero seguro.

Desde el punto de vista del método, no hay nada más decisivo que esa capacidad que el hombre tiene de comparar todo lo que existe con su corazón, tal que, incluso cuando se equivoca, aprende algo, ya que reconoce lo que corresponde y lo que no corresponde a lo que realmente desea y busca. Pero esto requiere una apertura del espíritu la cual tiene que educarse para comprender y utilizar adecuadamente el método ya que si bien es connatural al hombre no es algo espontáneo, esta apertura no se mantiene espontáneamente es necesario vivir intensamente la realidad.

La apertura a la realidad hay que educarla y para ello se requiere una atención hacia la realidad y una aceptación de la realidad. El prejuicio impide la atención porque afirma una idea preconcebida insensibilizando la capacidad de darse cuenta de la totalidad de los factores. Además, no permite acoger libremente y sin juzgar la realidad cerrando prematuramente la comprensión. La atención permite por tanto abrirse de par en par a la totalidad de los factores y la aceptación permite abrazar conscientemente lo que aparece delante de los sentidos y esto es fundamental para seguir un camino humano en el conocimiento.

Este documento es un pequeño aporte a la reflexión sobre el proceso de investigación científica desde un enfoque que pretende ser útil desde una visión comprensiva y humanista. Está conformado por dos secciones y la presente introducción. La sección que sigue a continuación abordan aspectos sobre aspecto epistemológicos sobre la ciencia y sobre el método de investigación. En la sección última, la más extensa, se focaliza ya sobre el cómo realizar un proceso de investigación; en cuanto a qué se investigara, cuál es su marco teórico, cuáles son los objetivos y el problema de estudio y el cómo se investigara en cuanto a la definición del tipo de estudio, los procedimientos, el alcance y los límites de la investigación.

---

<sup>1</sup> Académico de la Universidad de los Andes, adscrito al Instituto de Investigaciones Económicas Sociales de la Facultad de Administración y Ciencias Económicas y Sociales (IIES-FACES-ULA).  
<mailto:dramirez@ula.ve>

## 2. UNAS PRECISIONES EPISTEMOLÓGICAS

Para conocer se requiere tener conciencia de la coexistencia del ser, se conoce porque se coexiste con otros, se observa a sí mismo y a otros. No se puede pensar lo que no se conoce. Se investiga para conocer, ya que en el análisis del objeto, el sujeto puede encontrar el conocimiento y puede a su vez encontrarse a sí mismo (re-conocimiento). Esto resulta evidente cuando antes de hacer un juicio o establecer un concepto este se nutre de la experiencia cultivada. Esta experiencia parte de la premisa que "antes de la verdad existe lo verdadero"<sup>2</sup>. El conocer pone de relieve la indispensable coexistencia, copresencia y cooperación del sujeto que aprehende y de un objeto que es aprehendido.

En el DRAE<sup>3</sup>, se dice que la opinión es un juicio o un parecer que se forma de una cosa cuestionable, de hecho cuando se habla de "opinión pública" se refiere al juicio comúnmente aceptado y se dice que una persona se "casa con su opinión" cuando se aferra a su juicio propio. En la antigüedad se utilizaba una distinción entre la "opinión común" y la "opinión cultivada", para la "opinión común" se utilizaba la palabra "doxa" (δόξα) y para la "opinión cultivada" la palabra "episteme" (ἐπιστήμη), esta última se utiliza hoy en día para referirnos a la epistemología como al estudio de los fundamentos y métodos del conocimiento científico.

Cuando se habla de ciencia, se habla de ese conocimiento cultivado, de un saber científico. Para fines del documento se puede decir que "la ciencia es un modo de conocimiento que aspira a formular mediante lenguaje riguroso y apropiado —en lo posible, con auxilio del lenguaje matemático— leyes por medio de las cuales se rigen los fenómenos"<sup>4</sup>.

En las ciencias sociales, las leyes se entienden como regularidades de comportamiento probabilístico en cuanto a la relación causa efecto. La ley, en este sentido, indica que tal fenómeno ha de producirse según una ley determinada, pero no que el fenómeno en cuestión haya forzosamente de producirse. Estas leyes de comportamiento estadístico tienen en común con las leyes de las ciencias naturales los siguientes aspectos:

- a) Ser capaces de describir series de fenómenos
- b) Ser comprobables por medio de la observación de los hechos y de la experimentación
- c) Ser capaces de predecir acontecimientos futuros.

Así mismo como existen muchos modos de conocer en las diversas ciencias y disciplinas también existen muchas formas de comprobar y de predecir entre las diversas ciencias. La forma como se compruebe y como se predice depende del grado de desarrollo y formalidad que cada saber tiene. En el presente documento se trabaja sobre un modo de comprobar y predecir a través de técnicas cuantitativas.

Como ser humano se estudia lo que lo afecta, lo que le interesa, lo que lo motiva, de lo contrario no lo considera relevante o dentro de su campo de interés. La ciencia

---

<sup>2</sup> Étienne Gilson (197) **El realismo metódico**. Ediciones Encuentro, Madrid

<sup>3</sup> Diccionario de la Real Academia de la Lengua Española

<sup>4</sup> José Ferrater Mora (1980). **Diccionario de filosofía abreviado**. Editorial Sudamericana, Buenos Aires.

busca conocer para actuar en el objeto conocido con el fin de valorar, en la vida del hombre, los efectos que el objeto conocido provoca en ella, en su propia naturaleza.

El objetivo de toda investigación es conocer la verdad y para ello requiere formarse un juicio metódico sobre la realidad que se estudia. La capacidad de pensar es una verdad pero no el principio del conocer. El conocimiento parte del objeto, del dato bajo estudio que es fecundado por nuestra inteligencia. El objeto nos presenta así su naturaleza, sus elementos, sus matices, sus diferencias y semejanzas con otros objetos, datos y hechos. A este se le preguntara de una forma u otra, tal vez una forma de preguntar no dé respuesta alguna y por tanto se requerirá preguntar nuevamente de otra forma y desde otra perspectiva hasta hallar la forma, el método o el camino para que se obtenga alguna respuesta y se pueda entrar en el conocer. De aquí que sea el objeto el que define el método y no el método lo que define el objeto<sup>5</sup>.

Toda investigación se sustenta en valores éticos, entre los que destacada el amor a la verdad, la objetividad científica y un apego al realismo. El realismo exige que para conocer el objeto, el método no puede ser imaginado por él (o los) sujeto (s) sino que es impuesto por el objeto de investigación.

## **2.1. Sobre el Método de Investigación**

El conocer adecuadamente, es conocer con un método adecuado al objeto, pero este conocimiento esta limitado por nuestras formas de conocer. El "método científico" esta limitado en cuanto que para ser útil tiene que ser parcial y reduccionista.

La afirmación anterior puede sonar para algunos sorprendente pero baste por el momento un ejemplo. Si quisiera conocer una ciudad como Santiago, Buenos Aires o Caracas podría hacerlo de dos formas, una es recorriendo sus calles y otra consultado fotos y mapas. La primera me da una vivencia que no puede ser completamente transmitida, la segunda es replicable a otros. Pero si adicionalmente tuviese poco tiempo para visitar cada ciudad un mapa resultaría muy útil para focalizar mi tiempo en los puntos de interés, es decir me valgo de la experiencia acumulada por otros para optimizar mi conocimiento. Para que el mapa sea útil tiene que estar a una escala adecuada, reduciendo a lo mínimo lo esencial pero mostrando las principales rutas y sitios.

El mapa, en este caso, es una representación, un modelo parcial y reducido de la ciudad, no es la verdad pero muestra puntos verdaderos de la realidad en los cuales puedo focalizar mi atención y descubrir matices, hechos o datos que los otros no vieron. Nótese que tener toda la realidad implicaría tener un modelo a escala real uno a uno de la ciudad, pero no sería ni practico, ni útil. La ciencia pretende ser "mas" ciencia en la medida de que puede predecir acontecimientos o puede explicar mejor que el sentido común lo que acontece. Este afán de predecir y explicar nace del natural miedo humano de reducir la incertidumbre y valorar los riesgos. Puede que sea imposible impedir un huracán o un terremoto pero predecirlos permite reducir las perdidas de vidas y bienes.

---

<sup>5</sup> Luigi Giussani (1987) **El sentido religioso**. Ediciones Encuentro, Madrid

## 2.2. Paradigmas

En la ciencia se habla de paradigmas<sup>6</sup> que es como la "opinión mayoritaria" de los "expertos" de cómo estudiar el hecho científico, la historia de ciencia muestra que estos paradigmas cambian y son "validos" hasta cuando no son sostenibles por más tiempo. Estos paradigmas son el "modelo" de cómo hacer ciencia y no existe un consenso único por cuanto cada tipo de ciencia o disciplina requiere una forma particular de conocer.

Si se quiere conocer sobre matemáticas se utilizaría un método axiomático, deductivo. Si se quiere reacciones químicas, se utilizaría un laboratorio. Si se quiere saber cuantas personas pueden comprar un producto se realiza un estudio de mercado. Pero un error del paradigma es llevarlo al extremo de convertirlo en un sustituto de la realidad. En ese caso se cae en el "cientificismo" que es creer que la ciencia basta para satisfacer todas las necesidades de la inteligencia humana y que puede expresar todos los aspectos reales, morales e intelectuales del hombre. El error esta en creer que lo razonable esta en lo demostrable o lógico simplemente, cuando lo demostrable es sólo un aspecto de la razón en cuanto procedimiento que reproduce los pasos necesarios para construir, por ejemplo, un teorema a partir de unas premisas, así como la lógica ayuda a la demostración, por cuanto permite apuntar a la coherencia del procedimiento. Lo importante es saber que un enfoque o paradigma no sustituye a la realidad.

---

<sup>6</sup> Un paradigma es una concepción de cómo hacer un estudio científico, cuáles deben ser los problemas que deben abordarse, qué es lo que define el objeto, cuál es la naturaleza de los métodos y técnicas a utilizar, cuál es la información requerida y cómo se debe explicar, interpretar o comprender los resultados de la investigación.

**Cuadro 1**  
**Comparación de las investigaciones cuantitativas y cualitativas**

<b>Elementos</b>	<b>Tipo de Investigación</b>	
	<b>Cuantitativa</b>	<b>Cualitativa</b>
<b>Relación objeto-sujeto</b>	El investigador es el que decide, el investigado es tratado como objeto pasivo	Tanto el investigador como el investigado participan en todos los momentos de la investigación y viven la transformación
<b>Resultado/fin</b>	El producto es información. Es fundamentalmente para el conocimiento teórico. No lleva necesariamente a la realización de acciones concretas.	Involucra a los participantes en las acciones. Motiva a la acción
<b>Tipo de conocimiento</b>	Es un conocimiento puntual para un momento y un lugar dado	Estudia los fenómenos dentro de un contexto, dentro de un proceso en evolución
<b>Marco teórico</b>	Se define y construye al planificar la investigación	Se construye a través de los diferentes momentos de la investigación
<b>Selección de variables</b>	Pocas variables seleccionadas y definidas previamente	Las variables se van agregando y refinando en el estudio de campo
<b>Tipos de variables</b>	Enfatiza relaciones múltiples de variables de tipo cuantitativo o cuantificable. Variables nominales, de orden, de intervalo o de razón.	Énfasis en los aspectos subjetivos-cualitativos.
<b>Calidad de la información</b>	Enfatiza la confiabilidad	Enfatiza la validez
<b>Muestra</b>	Grande. Cuanto más mejor	Pequeña, adecuada al objeto
<b>Tipo de análisis principal</b>	Análisis de interdependencia y relaciones de causalidad	Análisis descriptivo e interpretativo

Fuente: Tomado y modificado de Pineda, Alvarado y Canales, 1994,13.

Estas notas metodológicas se inscribe dentro de un paradigma, que se considera útil pero no único, en el resumen del *Cuadro 1*<sup>7</sup> se muestran el enfoque cuantitativo, que es el que se abordará en el texto y el enfoque cualitativo que es igualmente válido para otros contextos. Estos dos ejemplos de enfoques muestran que cada uno aporta aspectos relevantes para explicar situaciones y construir indicadores que permitan describir y comprender la realidad. Mutuamente se ayudan y se pueden integrar en un proceso más amplio de estudio.

### 2.3. Investigación cuantitativa

La investigación cuantitativa se centra principalmente en tres aspectos; la investigación descriptiva, la relacional y la funcional o experimental: a) La *investigación descriptiva* es la etapa preparatoria del trabajo y permite ordenar el resultado de las observaciones de las conductas, de las características, factores y procedimientos observados, se trabajan sólo técnicas de análisis univariadas, b) la investigación *relacional* es más compleja respecto a la descriptiva y permite establecer comparaciones de variables entre grupos de estudio y control sin aplicar métodos estadísticos complejos más allá de las pruebas de inferencia de independencia y de contrastes de medias, y, por último, c) El análisis *funcional experimental* utiliza técnicas de análisis multivariado y permite construir modelos comparando grupos de control versus grupos bajo estudio.

### 2.4. Tipos de investigación

Toda investigación es tanto teórica como aplicada, ya sea cualitativa como cuantitativa y en esta última se destacan tres tipos o niveles en el proceso de investigación que son los de tipo exploratoria, descriptiva y explicativa

- a) *Exploratorias*: Son las investigaciones que pretenden darnos una visión general y sólo aproximada de los objetos de estudio.
- b) *Descriptivas*: Su preocupación primordial radica en describir algunas características fundamentales de conjuntos homogéneos de fenómenos.
- c) *Explicativas*: Son aquellos trabajos donde nuestra preocupación se centra en determinar los orígenes o las causas de un determinado conjunto de fenómenos. Su objetivo, por lo tanto, es conocer por qué suceden ciertos hechos, analizando las relaciones causales existentes o, al menos, las condiciones en que ellos se producen

## 3. SOBRE EL PROCESO DE INVESTIGACIÓN

¿Cuál es el camino? Es la primera y obvia pregunta cuando se quiere realizar un viaje, y como todo viaje hay un punto de partida y un punto de llegada pero también hay muchas rutas posibles. Lo mismo pasa en el método de investigación, existen diversas opiniones sobre qué estudiar, cómo se llama cada etapa, cuántas son, cuál es su orden, entre otros aspectos. Aquí se opta por un esquema simple pero no menos adecuado. Este se resume en la figura 1 que se explicará a continuación. En todo caso puede ser

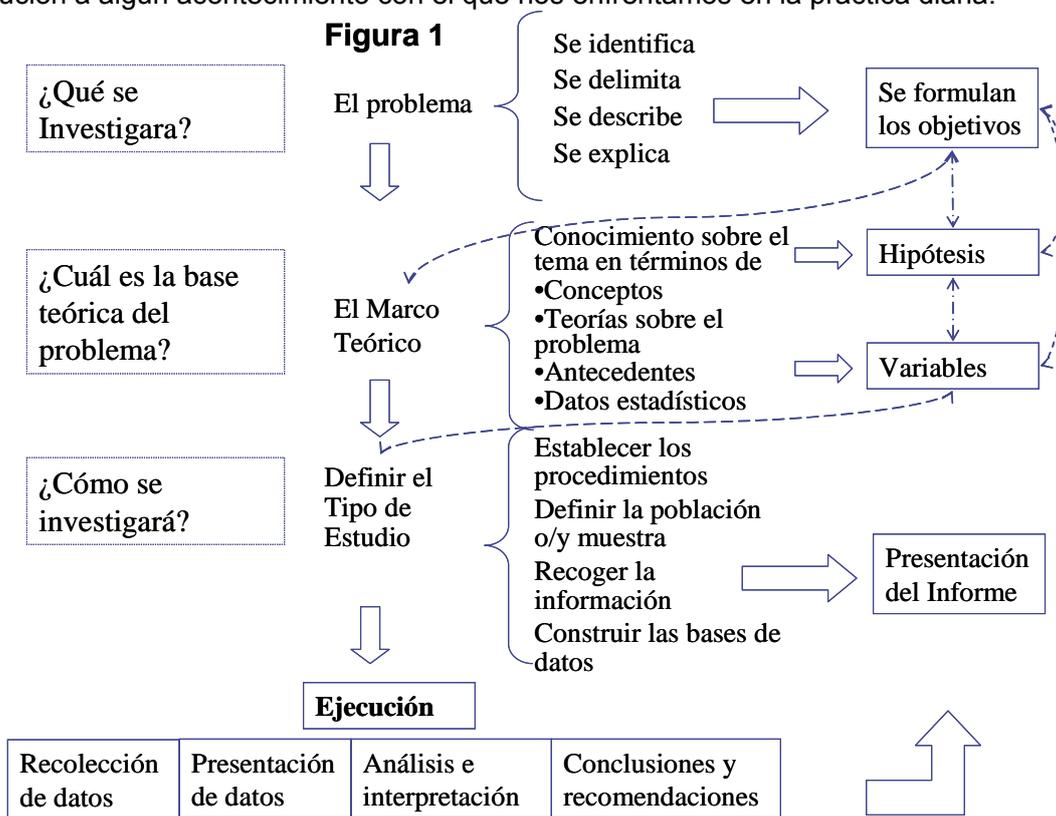
---

<sup>7</sup> Tomado y modificado de Pineda, E. B., E. L. de Alvarado y †F. H. de Canales (1994) **Metodología de la investigación**. Organización Panamericana de Salud. Washington, D. C. p. 13.

contrastado con otros textos más formales pero en la comparación notaran que se toma lo esencial del proceso. Nótese que aun cuando existe cierto orden lógico y cronológico del proceso puede que excita un proceso de retroalimentación entre las partes como la retroalimentación del marco teórico en la reformulación de los objetivos y a su vez de estos a las hipótesis y a las variables a incluir en el estudio.

### 3.1. ¿Qué se investigara?

En términos simples un problema es una cuestión que se quiere aclarar, en las matemáticas un problema es una proposición dirigida a averiguar el modo de obtener un resultado cuando ciertos datos son conocidos. En el campo de la investigación se percibe como un vacío o la ausencia de una respuesta ante una realidad dada y este se da porque ignoramos como ocurren ciertos fenómenos, o cuando no podemos explicarnos por qué ocurren. Esto refleja las dos grandes fuentes de temas de investigación, la primera surge porque se encuentran vacíos o inconsistencia en el conocimiento acumulado en el área de estudio y la segunda es porque no se tiene una respuesta o solución a algún acontecimiento con el que nos enfrentamos en la práctica diaria.



#### 3.1.1. El planteamiento del problema

La identificación de los hechos, fenómenos o acontecimientos que se desean conocer, delimitando describiendo y explicando la situación que se quiere estudiar. Se parte por reconocer él o los hechos que involucra el fenómeno bajo estudio. Termina

definiendo el problema que se quiere estudiar (objetivo central o general) y el cómo abordarlo (objetivos específicos) por etapas sucesivas o sub-problemas.

Para el principiante definir un problema representa una mayor dificultad que para quien ya está investigando en un campo o que tiene una mayor práctica sobre la cual reflexiona, o trabaja con mayor frecuencia, o que lee sistemáticamente, por ello se aconseja trabajar sobre lo más conocido. Como ayuda y sugerencia se plantea algunos criterios para seleccionar un tema de investigación los cuales se pueden resumir en las siguientes líneas:

1. *Escoger una temática conocida.* Es conveniente concentrarse en aquellas áreas del conocimiento sobre las que ya se posee una información más o menos considerable y postergar para mejor ocasión el abordaje de temas sobre los que no se conoce lo suficiente
2. *Seleccionar un tema bien concreto y accesible:* Es preferible definir un tema concreto, bien delimitado y preciso, sobre el que pueda encontrarse suficiente bibliografía y sea factible obtener suficientes datos de interés.
3. *Buscar áreas de trabajo en las que pueda contarse con una ayuda efectiva.* Un dicho dice que la rueda ya está inventada y lo que toca es saber usarla, ningún investigador puede crear todo el conocimiento se necesita de otros y de sus experiencias para llegar un poco más adelante en el conocimiento. Por ello siempre es recomendable contar con apoyo externo cuando se comienza a investigar y resulta útil escoger temas en los que ya se esté trabajando y en los cuales pueda contarse con el asesoramiento y la orientación de profesores e investigadores más versados en la materia
4. *Buscar un problema de investigación que resulte de real interés.* El tener una genuina curiosidad para estudiar el tema seleccionado provee las energías y la motivación para desplegar los mejores esfuerzos en la superación de los inevitables inconvenientes que siempre se ha de enfrentar durante el proceso. La práctica de la investigación nos suele enfrentar a dificultades que generan ansiedades, ante las tareas rutinarias que se pueden volver tediosas, o ante el esfuerzo que, en general, resulta bastante intenso y sostenido. Si no se tiene una pasión y un interés sobre el tema será difícil poder alcanzar la meta del informe final

Muchas veces se destaca como una debilidad en la investigación que el investigador es un sujeto con una percepción y una posición sobre la realidad. Esto más que una debilidad es un dato que debe ser incorporado y utilizado ya que toda investigación científica debe considerar todos los datos o factores de la realidad. Para explicarlo tomare un ejemplo de Giussani<sup>8</sup>. Si llamamos *r* (o "razón") o capacidad de conocer del sujeto y llamamos *v* (o "valor") a la realidad que se quiere conocer una vez que entran en el campo de interés humano y llamamos *s* (o "sentimientos") la presencia intermediaria.

$$r \rightarrow s \rightarrow v$$

La respuesta de algunos es eliminar la "s" con lo cual se llega a la paradoja de cuanto algo más me interesa conocer menos me es posible ya que mis sentimientos se interponen, en segundo lugar se cae en el error de eliminar un factor de la realidad que está en juego. Giussani responde que "la cuestión no es eliminar el sentimiento, sino situar al sentimiento en su justo lugar". El centro del problema es tener una postura una

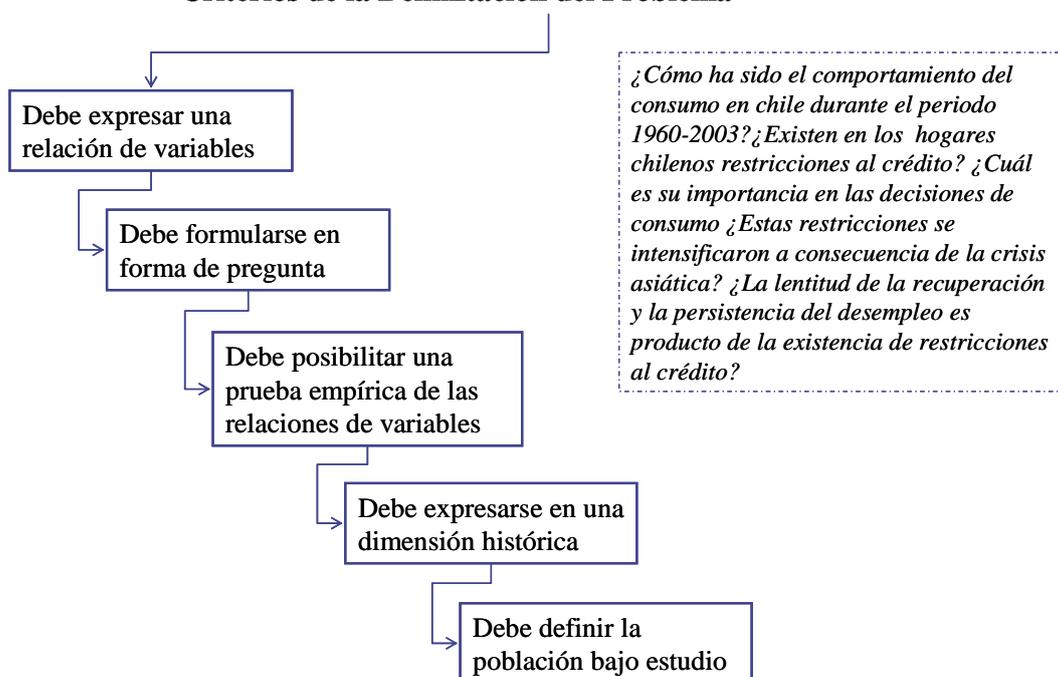
---

<sup>8</sup> Giussani, Ob. Cit., pp. 43, 45 y 46

actitud adecuada, un sentimiento apropiado, una moralidad y esta se sustenta en "que el amor a la verdad del objeto sea siempre mayor que el apego a las opiniones que uno tiene de antemano sobre él" y se puede concluir que "la suprema pobreza de espíritu es la pobreza ante la verdad".

**Figura 2**

**Criterios de la Delimitación del Problema**



Fuente: Tomado y modificado de Pineda, Alvarado y Canales, 1994,43.

En la figura 2<sup>9</sup> se presenta un esquema sobre los criterios que delimitan el planteamiento del problema los cuales se describen a continuación:

1. *Debe expresar una relación de variables*: La claridad de un problema se encuentra en la medida que se conocen sus dimensiones o características fundamentales. Las variables son características o factores de la población bajo estudio que se pueden expresar en unidades o elecciones, estas se relacionan entre sí y conforman estructuras que son las que generan situaciones problemáticas o no. El describir el problema en términos de sus variables permite definir los objetivos y construir el marco teórico.
2. *Debe expresarse en forma de pregunta*: El plantear el problema en forma de pregunta exige responder y aclarar más lo que se va a investigar, puesto que esto ayuda a sistematizar las ideas y revisar diversas posiciones que existan al respecto con lo que nos introduciremos en la fase siguiente de elaboración del marco teórico. Otra alternativa es describir la situación con el mayor número de aristas y perspectivas posibles de la situación que se pretende resolver o descubrir.

<sup>9</sup> Tomado y modificado a partir de Pineda, Alvarado y Canales, 1994,43.

3. *Debe ser posible de verificar una comprobación empírica de las variables:* El propósito de la investigación es comprobar o rechazar algo. Plantear un problema es afinar su estructura formalmente. Se debe describir el problema en términos concretos, explícitos y específicos, de manera que los argumentos puedan ser investigados por medio de los procedimientos científicos. Un problema correctamente planteado está parcialmente resuelto, a mayor exactitud corresponden más posibilidades de obtener una solución satisfactoria.
4. *Debe expresarse en una dimensión espacio-temporal concreta:* La definición espacio temporal de la investigación resulta práctico y delimita el alcance de la investigación, ya que no se puede estudiar todo ni tampoco todo el tiempo. Además las variables involucradas no se presenta siempre en el mismo estado en dos lugares y momentos distintos<sup>10</sup>. Además de definir cuando y como permite evaluar la factibilidad de realizar el estudio en el tiempo disponible y con los recursos que se cuentan.
5. *Debe delimitarse claramente la población bajo estudio:* El investigador debe ser capaz no solo de describir y entender el problema sino también de verbalizarlo en forma clara, precisa y accesible. En algunas ocasiones sabe lo que desea hacer pero no sabe o puede comunicarlo a los demás y es necesario que realice un esfuerzo por traducir su pensamiento a términos que el mismo entienda y acepte para después poder comunicarlo a los demás. El problema debe estar formulado claramente; describir los hechos, situaciones, participantes, características del fenómeno, lugares, fechas, conflictos, dramas, situaciones difíciles, desenlaces, etc... que permita definir claramente porque lo considera un problema y para quien –entendiendo por problema, la problemática planteada o el fenómeno estudiado-

### 3.1.2. Objetivos de Investigación

Una vez seleccionado un problema de interés y que en principio resulte factible se hace necesario preguntar ¿cuál es el sentido de estudiar el problema? y ¿qué conocimientos se puede obtener de esta investigación?, en esta fase inicial del trabajo, los objetivos básicos que persigue la investigación, debe ser definidos con rigurosidad y en forma explícita.

El resultado final de esta etapa se encuentra en la formulación del objetivo general que es como se expresa el asunto de interés para el investigador. La palabra objetivo proviene de ob-jectum, que significa "*a donde se dirigen nuestras acciones*" y se entiende como el término o fin de los actos o como el fin o intento a que se dirige o encamina una acción u operación. Los objetivos son enunciados escritos sobre resultados a ser alcanzados en un periodo determinado, así como se muestra descriptivamente en la figura 3<sup>11</sup>.

<sup>10</sup> Sirva de ejemplo esta idea simple, en un sistema de dos variables podemos tener 4 estados de relaciones posibles, por ejemplo no vincularse, vincularse simultáneamente vincularse de un lado hacia el otro y viceversa, para estimar los estados posibles de un sistema se representa por una relación  $2^{(N*(N-1))}$ ; si N es igual a 2 componentes o variables el resultado sería 4, pero si los componentes fueran 5 el número de estados posibles sería de más de un millón de posibilidades (exactamente de 1.048.576)

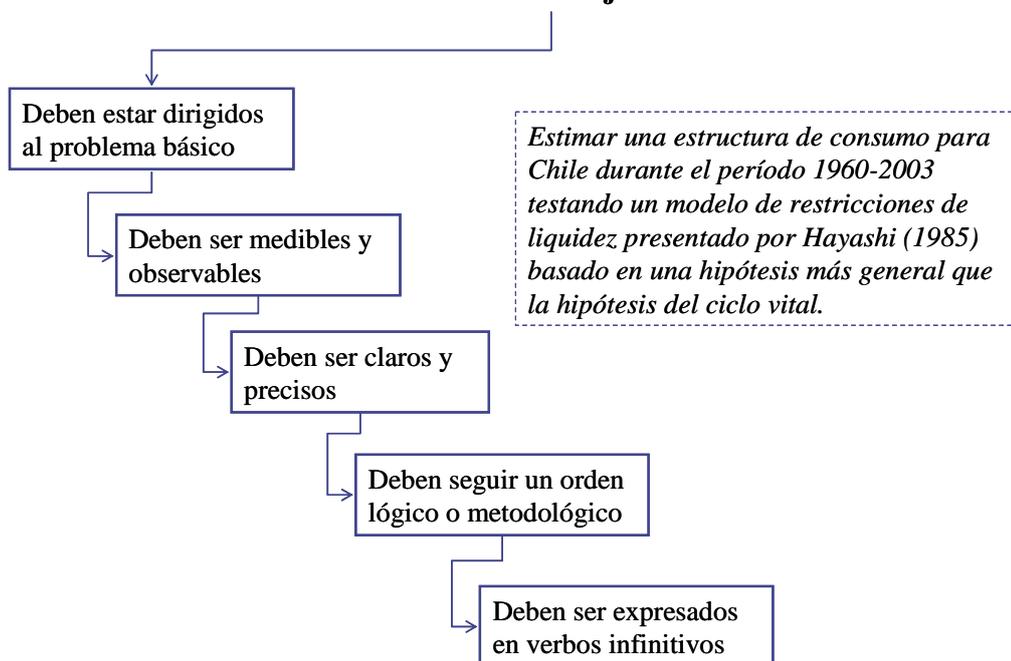
<sup>11</sup> Tomado y modificado de Pineda, Alvarado y Canales, 1994, 48.

El objetivo formulado por el investigador debe ser uno y solo uno, más de un objetivo general expresa que el propósito de la investigación no está claramente definido por tanto se debe revisar de nuevo la formulación del problema o propósito de investigación. Este propósito debe responder al asunto en cuestión o problema básico de interés. Pero se refinan en la medida en que se profundiza el marco teórico conceptual del estudio

Los objetivos de investigación tienen como propósito: a) orientar las demás fases del proceso de investigación, b) determinar los límites y la amplitud del estudio, c) Permitir definir las etapas que requiere el estudio y, por último, d) situar el estudio dentro de un contexto más general.

**Figura 3**

**Criterios de la Definición de Objetivos**



Fuente: Tomado y Modificado de Pineda, Alvarado y Canales, 1994,48.

Para ello deben estar dirigidos al problema planteado con el fin de dar un propósito a la investigación, y estos deben ser evaluables o medibles y observables en su cumplimiento o no.

- a) Los objetivos se relacionan directamente con el problema a estudiar y siguen un orden lógico, se redactan vinculando las variables relevantes y se expresan con verbos en infinitivo, tales como: identificar, determinar, establecer, distinguir, cuantificar, medir.
- b) Los objetivos deben ser redactados de forma operacional que permitan ser medibles y observables y así poder ser evaluados o verificados.
- c) Cada objetivo debe estar enfocado a un sólo aspecto del problema y deben ser claros y precisos

### 3.2. ¿Cuál es la base teórica del problema?

Definido el problema de estudio es necesario recopilar los antecedentes, tanto estadísticos, como bibliográficos, que permitan construir un marco comprensivo del fenómeno y permita definir los factores que describen las relaciones y ayuden al análisis, comprensión y explicación del problema. La validez interna y externa de una investigación se demuestra en las teorías que la apoyan y, en esa medida, los resultados pueden generalizarse. El marco teórico de la investigación considera:

- Los conceptos explícitos e implícitos del problema
- Las metodologías utilizadas para analizar este tipo de problemas o similares.
- Las relaciones teóricas y los conceptos adoptados.
- Elaborar los análisis teóricos de los cuales se desprenden las hipótesis.
- Las implicaciones teóricas del problema.-

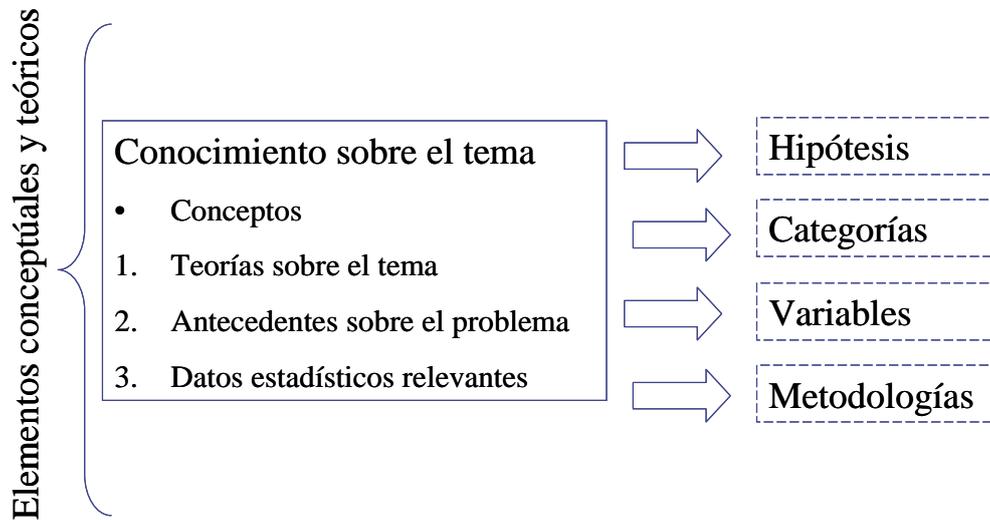
El la revisión bibliográfica y de antecedentes, conocida como marco teórico se elabora una síntesis del contexto general (local, nacional y mundial) en la cual se ubica el tema de la propuesta, se establece el estado actual del conocimiento del problema, las brechas que existen y el vacío que se quiere llenar con el proyecto; el porqué y el cómo la investigación propuesta, se va a diferenciar de las investigaciones previas, y cómo contribuirá, con probabilidades de éxito, a la solución o comprensión del problema planteado. Las funciones principales del marco teórico son las siguientes:

- Orientar el conocimiento del tipo de investigación
- Ampliar el horizonte de estudio, permitiendo expresar distintos puntos de vista
- Conducir al establecimiento de hipótesis
- Inspirar líneas de investigación
- Prever la manera de interpretar los datos provenientes de la investigación

#### 3.2.1. Formulación del Problema

La formulación del problema, la construcción del marco conceptual y teórico y el planteamiento del problema son procesos que se distinguen lógicamente pero en términos cronológicos se trabajan casi en paralelo, como se trata de ejemplificar en la figura 1 y en la figura 4, pero se retroalimentan y se refuerzan mutuamente y a su vez deben mantener una coherencia lógica o congruencia entre las partes.

**Figura 4**



Fuente: Elaboración propia

### 3.2.2. Construcción del Marco Teórico

Para la construcción del marco teórico debe plantearse de forma explícita todos los elementos del fenómeno en el proceso de análisis. Estos se pueden describir a través de los siguientes pasos:

1. *Identificar los elementos que fundamenta el Tema:* Consiste en identificar los elementos teóricos para fundamentar el problema. De la revisión de antecedentes se obtienen los resultados en términos de factores y aspectos para fundamentar el problema y los resultados de las diferentes teorías, investigaciones y datos estadísticos que a juicio del investigador son pertinentes y le permita manejar con experticia el código de comunicación del campo de trabajo.
2. *Selección de Variables principales:* Seleccionar las variables principales o elementos más relevantes para el estudio del problema y de la revisión de la literatura adecuada. La variable central o variable dependiente se refiere básicamente al problema y las variables explicativas o causales son las que explican y causan la respuesta de la variable dependiente y son conocidas como variables independientes, los parámetros que vinculan a ambas son conocidos como parámetros estimados de comportamientos y en el análisis de regresión como coeficientes de regresión o regresores.
3. *Identificar las relaciones y enunciar las hipótesis:* Identificar las relaciones entre las variables y enunciar las hipótesis centrales y o complementarios con sus posibles pruebas de contrastes a que deben ser sometidas para su validez o falsificación.
4. *Diseño Experimental:* Construir el esquema de las relaciones entre variables o constructos teóricos señalando sus vinculaciones principales y sus indicadores empíricos o variables operacionales, esto ayuda a tener una visión del conjunto.
5. *Elaborar Propuesta:* Elaborar el marco conceptual del estudio que establezca los antecedentes, proposiciones y metodologías específicos para el problema.

### 3.2.3. Partes del Marco Teórico

El Marco Teórico contiene tres partes claves:

1. *El marco teórico propiamente dicho*: Es un conjunto de proposiciones referidas al problema de investigación tomadas de una o más teorías existentes sobre el campo donde éste se ubica, con las modificaciones que el investigador esté en condiciones o capacidad de introducirles, contiene los antecedentes que se tienen sobre el problema suelen tener una mayor consistencia lógica, de tal modo que el problema resulta como derivado o deducido de ese conjunto conceptual.
2. *El marco referencial*: Está constituido por el conjunto de conocimientos que otros estudiosos han logrado sobre el tema o el problema de investigación que se ha propuesto un investigador. Tanto este marco como los otros proporcionan un contexto de referencia del problema a investigar.
3. *El marco conceptual*: Es, como lo indica su nombre, una elaboración conceptual del contexto en el cual se considera el problema. Está compuesto de referencias a sucesos y situaciones pertinentes, a resultados de investigación —incluye, por tanto, un marco de antecedentes—, definiciones, supuestos, etc. Se podría decir que este marco es una especie de constructo teórico, sin grandes pretensiones de consistencia lógica entre las proposiciones que la componen, aun si utiliza conceptos de alguna teoría existente. Además de establecer el tipo de pruebas de verificación para comprobar la certeza de las hipótesis planteadas.

#### 3.2.4. Formulación de Hipótesis<sup>12</sup>

Las hipótesis, con la selección de las variables teóricas y operacionales, son los principales productos de la revisión bibliográfica y de antecedentes en el proceso de elaboración del marco teórico. Durante el proceso se debe diseñar la prueba de contraste, especificando qué metodología se utilizará para encontrar la respuesta al problema, en quiénes se estudiarán y qué métodos se aplicaran para la obtención y análisis de los datos y resultados.

Las hipótesis es una suposición o proposición que establece la existencia de una relación entre dos o más variables expresadas como hechos, fenómenos o factores o entidades que deben ser sometida a prueba para ser aceptada como válida o rechazada como falsa y su función consiste en delimitar el problema que se va a investigar. En este último sentido constituye un enlace o puente entre la teoría y la investigación.

Las hipótesis se definen en términos de: a) su unidad de análisis que pueden ser individuos, familia o instituciones, b) las variables que la conforman, entendidas como las características, factores o propiedades de las unidades de análisis y pueden ser variables del tipo ordinal, ordinal de intervalo o de razón, y, c) por los elementos lógicos, que relacionan las unidades de análisis con las variables y a estas entre sí en términos de: mayor(>), menor(<) igual, conjunción ( $\wedge$ ) exclusión ( $\vee$ ) negación ( $\sim$ ); entre otros signos<sup>13</sup>.

<sup>12</sup> Sólo en los casos de investigaciones explicativas es necesario formular claramente cuáles son las hipótesis de la investigación. En las investigaciones descriptivas y, con más razón, en las exploratorias, es posible omitir las hipótesis, ya sea porque éstas son tan amplias y poco definidas que dicen muy poco a quien lee el informe de investigación, o porque no es posible o necesario verificarlas.

<sup>13</sup> Existen muchos signos de relaciones binaria para el caso solo se han señalados los básicos de la lógica simbólica, en economía se utilizan los símbolos de preferencia y en la teoría de juegos los símbolos de precedencia para construir árboles de juego.

### 3.2.5. Deducción e Inducción

Las hipótesis pueden ser formuladas a partir de un proceso deductivo o inductivo. Las *hipótesis deductivas* tienen como punto inicial una teoría o una proposición general que se aplica a situaciones particulares a través de un razonamiento deductivo (de lo general a lo particular) se postulan principios teóricos a ser comprobados. Las *hipótesis inductivas* parten de las observaciones particulares de los hechos y donde se detecta un cierto patrón recurrente que aparentemente representa una tendencia o fenómeno recurrente que expresa una asociación particular y que posteriormente, a partir de estas observaciones parciales, se quiere comprobar si expresa una ley general.

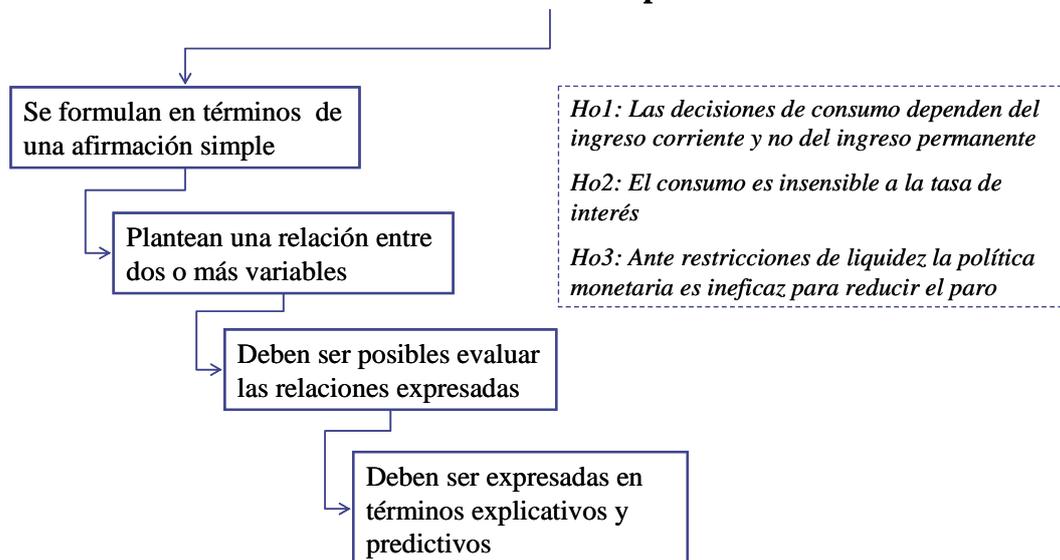
### 3.2.6. Criterios para la Formulación de Hipótesis

Los criterios para la formulación de una hipótesis tienen una gran correspondencia con los criterios para el planteamiento de un problema y con los criterios para formular los objetivos para ello véase y compare las figuras 2, 3 y 5.

- a) El enunciado o proposición debe estar expresado de manera positiva y simple, por ejemplo, “bajo restricciones de liquidez el consumo es insensible a la tasa de interés” Esto quiere decir que el parámetro o pendiente que relaciona la tasa de interés como variable explicativa no afecta al consumo o que estadísticamente su valor es cero.
- b) Debe expresar una relación entre las variables, en el ejemplo se tiene una relación entre la variable dependiente “consumo” y la variable explicativa “tasa de interés”.
- c) La relación entre las variables planteadas en la hipótesis pueden ser sometidas a pruebas de contraste sobre la validez de la afirmación, para el ejemplo se pueden hacer pruebas de correlación, y de significancia de la correlación, también se pueden hacer pruebas de independencia entre otras.
- d) Por último deben ser capaces de predecir un comportamiento en este caso por ejemplo; cambios en la tasa de interés no van a modificar el patrón de consumo de los hogares que presentan restricciones de liquidez.

**Figura 5**

**Criterios de la Formulación de Hipótesis**



Fuente: Tomado y Modificado de Pineda, Alvarado y Canales, 1994,66.

Las hipótesis, al igual que los objetivos, tiene como requisitos que al ser formuladas deben referirse a un sólo ámbito de la realidad, deben ser claras y precisas en sus términos de referencia y en su planteamiento debe estar implícito la técnica para probarla por ello se recomienda que

- a) Los términos que se empleen deben ser claros y concretos para poder definirlos de manera operacional, a fin de que cualquier investigador que quiera replicar la investigación, pueda hacerlo. Una hipótesis sin referencia empírica constituye un juicio de valor. Si una hipótesis no puede ser sometida a verificación empírica, desde el punto de vista científico no tiene validez.
- b) Las hipótesis deben ser objetivas y no llevar algún juicio de valor; es decir, no debe definirse el fenómeno con adjetivos tales como "mejor" o "peor", sino solamente tal y como pensamos que sucede en la realidad.
- c) Las hipótesis deben ser específicas, no sólo en cuanto al problema, sino a los indicadores que se van a emplear para medir las variables que estamos estudiando.
- d) Las hipótesis deben estar relacionadas con los recursos y las técnicas disponibles. Esto quiere decir que cuando el investigador formule su hipótesis debe saber si los recursos que posee son adecuados para la comprobación de la misma.
- e) La hipótesis debe estar directamente relacionada con el marco teórico de la investigación y derivarse de él.

**3.2.7. Hipótesis y nivel de error**

Generalmente las hipótesis, al ser enunciadas para su comprobación o falsificación estadística a través de pruebas inferenciales, se clasifican en hipótesis nula y alternativa. Estas se postulan cuando se trata de discernir si existen diferencias

significativas o no entre las estimaciones obtenidas y los parámetros de la población o, entre estimaciones obtenidas en diferentes muestras.

La hipótesis nula se designa con el símbolo  $H_0$ . La hipótesis nula es aquella que nos dice que no existen diferencias significativas entre los grupos. Una hipótesis nula es importante por las siguientes razones: *primero* porque es una hipótesis que se acepta o se rechaza según sea el resultado de la investigación y en *segundo* lugar por el hecho de contar con una hipótesis nula ayuda a determinar si existe una diferencia significativa entre los grupos y si no se debió al azar. En la mayoría de los casos se establece con el propósito explícito de ser rechazada, en este caso se espera que exista diferencias en el resultado y la conclusión que se espera alcanzar es que se “*rechaza la hipótesis nula*”, en caso contrario, de que los datos disponibles no permitan encontrar evidencia para rechazar la hipótesis nula, la conclusión sería “*no se puede rechazar la hipótesis nula*”. La hipótesis nula es una hipótesis simple ya que se compara con un valor puntual y no asume afirmaciones del tipo “mayor que” o “menor que” siempre se formula en términos de “ser igual a”.

La hipótesis alternativa se designa con el símbolo  $H_a$ . En ella se postula que las estimaciones o valores observados en el universo, en la muestra o en una de las variables en estudio no son equivalentes a los encontrados en otro universo, u otra muestra o variable bajo análisis. Lo que significa que sí existen diferencias significativas entre los valores estimados. En esta circunstancia el valor obtenido puede ser mayor (>) o menor (<) al valor comparado como umbral. Por ello se dice que la hipótesis alternativa es una hipótesis compleja ya que implica mas de un valor o resultado posible.

**Figura 6**

- $H_0$  : se conocerá como la hipótesis nula. Esta es una hipótesis simple, la cual se quiere probar o rechazar.
  - Error Tipo I =  $P(\text{rechazar } H_0 \mid H_0 \text{ es cierto})$
- $H_a$ : se conocerá como la hipótesis alternativa. Esta es una hipótesis compuesta, lo cual es casi la negación de  $H_1$ .
  - Error Tipo II =  $P(\text{aceptar } H_0 \mid H_a \text{ es cierto})$

		<i>Realidad</i>	
		Ho: Cierta	Ha: Falsa
<i>Decisión estadística</i>	Ho: Aceptar	(1- $\alpha$ ) Confianza	$\beta$ Error Tipo II
	Ha: No aceptar	$\alpha$ Error Tipo I	(1- $\beta$ ) Potencia

Fuente: Tomado y modificada a partir de Anderson, Black, Tathan y Hair (1999) **Análisis Multivariante**. Prentice Hall, Madrid.

Para interpretar las pruebas de inferencias estadísticas el investigador especifica los niveles aceptables de error estadístico. Los errores estadísticos se denominan comúnmente con los nombres de “Error Tipo I” y “Error Tipo II”. El nivel del error Tipo I, también se conoce como alfa<sup>14</sup> ( $\alpha$ ) y se define como la probabilidad de rechazar la hipótesis nula ( $H_0$ ) cuando es cierta (“un positivo falso”), o dicho de otra forma, si partimos de la presunción de inocencia del inculpado entonces la hipótesis nula sería “ $H_0$ : Es inocente”, entonces el error tipo I sería “declarar inocente a un culpable”. El investigador define el nivel de error Tipo I, al establecer los márgenes aceptables de error, por convención se asume un error del 5% como nivel típico aceptable. Una probabilidad más interesante es  $(1-\alpha)$  que se denomina como nivel de **confianza** de la prueba de contraste con que se rechaza la hipótesis nula. En la figura 6<sup>15</sup>, se relaciona el tipo de error con los tipos de hipótesis de investigación.

Al definir el error Tipo I, el investigador también define el nivel de error Tipo II que es un error asociado y se conoce como error beta ( $\beta$ )<sup>16</sup>. El error Tipo II es la probabilidad de fallar en rechazar la hipótesis nula cuando realmente es falsa. En otras palabras, siguiendo el ejemplo anterior, sería “declarar inocente a un culpable”. La probabilidad  $(1-\beta)$ , se denomina **potencia** de la prueba de contraste. La potencia es la probabilidad de rechazar correctamente la hipótesis nula cuando debe ser rechazada.<sup>17</sup> Los errores Tipo I y Tipo II se encuentran inversamente relacionados, y a medida que el error del Tipo I se reduce el error de Tipo II aumenta, ya que al reducir el error Tipo I también se reduce el poder de la prueba estadística por ello el analista tiene que encontrar un equilibrio entre el nivel del alfa y el nivel de potencia.<sup>18</sup>

### 3.2.8. El diseño metodológico

El diseño metodológico del estudio se describe en el esquema anexo en la figura 7, como se muestra en el esquema, definido los objetivos de investigación, se debe decidir el cómo se recopila la información, cómo se operacionaliza, en términos del cuestionario, cómo se evalúa el instrumento y luego, cómo se levanta la información. El diseño metodológico es la base para planificar todas las actividades que demanda el proyecto y para determinar los recursos humanos y financieros requeridos. Una metodología vaga o imprecisa no brinda elementos para evaluar la pertinencia de los recursos empleados.

El diseño metodológico debe mostrar, en forma organizada, clara y precisa, cómo se alcanzarán cada uno de los objetivos específicos propuestos. La metodología debe reflejar la estructura lógica y el rigor científico del proceso de investigación desde la elección de un enfoque metodológico específico (preguntas con hipótesis fundamentadas correspondientes, diseños de la muestra o experimentales) hasta la forma como se van a analizar, interpretar y presentar los resultados.

---

<sup>14</sup> Anderson, Black, Tathan y Hair (1999) **Análisis Multivariante**. Prentice Hall, Madrid

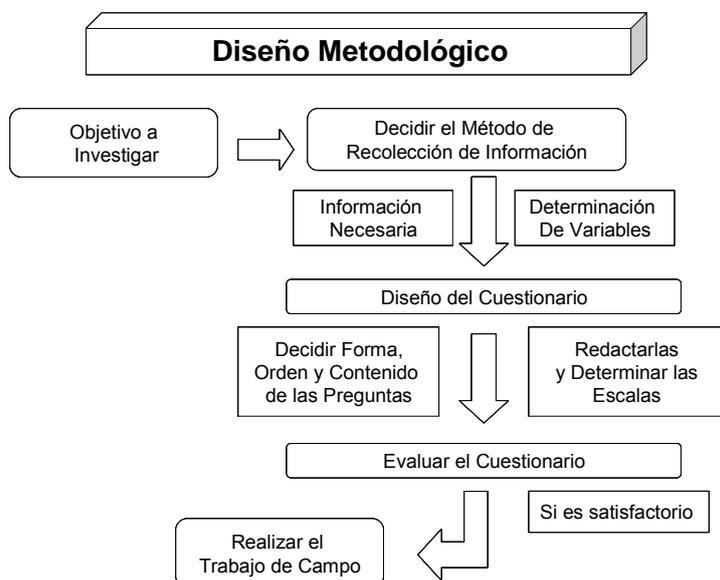
<sup>15</sup> La figura 6 es tomada y modificada de: Anderson, Black, Hair y Tathan; Ob. Cit..

<sup>16</sup> Anderson, Black, Tathan y Hair (1999) Ob. Cit.

<sup>17</sup> Ob. Cit.

<sup>18</sup> Para más referencia ver; Anderson et. al. Ob. Cit.

**Figura 7**



Fuente: Elaboración propia.

En el diseño metodológico deben detallarse, los procedimientos, técnicas, actividades y estrategias metodológicas requeridas para la investigación. Deberá indicarse, por tanto, el proceso a seguir en la recolección de la información, así como en la organización, sistematización y análisis de los datos.

### 3.2.9. Diseño Metodológico de la Muestra<sup>19</sup>

La inferencia estadística permite conocer las características más relevantes de una población a partir de observaciones parciales o muestras de la misma, permitiendo la inducción de los resultados estadísticos bajo el supuesto que puede expandirse los resultados a la población en estudio. La muestra estadística es una parte de la población con el fin de investigar, a partir del conocimiento de sus características particulares, las propiedades de toda la población. La población se entiende como un conjunto finito o infinito de personas o elementos que presentan características comunes.

La muestra debe ser representativa del universo de la cual se obtuvo y debe ser aleatoria en cuanto a la manera de elegirla. Para que la muestra sea aceptable, desde el punto de vista del análisis estadístico, deberá reunir las siguientes características:

- Deberá seleccionarse de manera que sea representativa de la población de la cual se obtuvo.
- Deberá ser confiable tal que permita obtener estimaciones de las características de la población y proporcionar medidas de la precisión del error de estimación.
- Deberá ser la que provea mayor información con el menor costo. Aun cuando esta última no es indispensable constituye un criterio deseable.

<sup>19</sup> Esta sección se refiere al caso de que se tenga que realizar un diseño muestral para la investigación, en estudios de series de tiempo normalmente no se tiene este tipo de situaciones.

La muestra óptima es aquella que reúne las exigencias de eficiencia, representatividad, seguridad y flexibilidad. La muestra más eficiente es la que provee mayor información por el menor costo. Siendo lo bastante pequeña como para evitar gastos superfluos y lo suficientemente grande como para que el error de la muestra sea aceptable. La muestra debe ser representativa, en cuanto debe permitir obtener estimaciones de los valores de la población o universo a partir de las medidas obtenidas. Debe ser segura porque permite calcular medidas de la confiabilidad de las estimaciones. Por último, debe permitir interpretar estas estimaciones de la manera más precisa, a través de las pruebas estadísticas de hipótesis acerca del universo.

La utilización o selección de una muestra para un estudio es aconsejable en los siguientes casos:

- Cuando la población sea muy grande y resulte muy costoso (para los recursos disponibles) la enumeración completa de cada uno de los elementos que compone la población. Por ejemplo un estudio sobre intención del voto.
- Cuando la población es lo suficientemente homogénea que permite que una muestra sea una imagen aceptable y representativa de la población. Por ejemplo un estudio sobre el rendimiento promedio de un cultivo de papa en Mucuchies.
- Cuando en el proceso de investigación de la característica bajo estudio produzca un daño irreversible en el elemento o individuo seleccionado. Por ejemplo un estudio para determinar la antigüedad de una pieza arqueológica sometida a Carbono 14.

Adicionalmente a las anteriores razones se pueden añadir otras limitaciones y ventajas que se resumen en el cuadro que se muestra en el Anexo 3.

El muestreo aleatorio simple permite seleccionar una muestra sin conocer previamente las características del universo. Permite, además, comprobar fácilmente la exactitud de las estimaciones ya que los errores del muestreo están sometidos a las leyes estadísticas y por último, permite obtener la información estadísticamente más segura por un costo mínimo. Por tanto, las unidades seleccionadas representan mejor el universo en sus características principales y a medida que aumenta el tamaño de la muestra, se incrementa la representatividad de la misma.

Existen dos tipos de muestreo aleatorio simple el con reemplazamiento y el sin reemplazamiento. El muestreo aleatorio simple con reemplazamiento otorga la misma probabilidad de selección a todos los elementos de la población y coincide con el muestreo de poblaciones infinitas ya que al ser devuelto cada elemento extraído a la población (luego de ser anotada sus características) la población es inagotable y el resultado de la extracción es independiente de las extracciones anteriores. El muestreo aleatorio simple sin reemplazamiento, al igual que el método anterior tiene la misma probabilidad de ser extraída, pero la población es finita, por lo tanto, la probabilidad de que salga un elemento dependerá de los que fueron separados anteriormente para formar parte de la muestra, y dejaron de ser seleccionables.

El muestreo estratificado aleatorio es un tipo de muestreo probabilístico que tiene por objeto mejorar las estimaciones del muestreo aleatorio simple. A través del agrupamiento previo de los elementos más parecidos entre sí, se divide la población en

subpoblaciones o estratos, dentro de cada uno de los cuales se hace una selección aleatoria simple que puede ser con reemplazamiento o sin reemplazamiento.

El reparto del tamaño de la muestra entre los estratos o subpoblaciones suele llamarse afijación. Esta afijación puede ser de tres tipos.

- Afijación uniforme. Consiste en repartir por igual la muestra entre todos los estratos
- Afijación proporcional. Consiste en distribuir la muestra proporcionalmente al número de elementos de cada estrato
- Afijación óptima. Consiste en hacer el reparto proporcional al número de elementos y a la desviación estándar de cada estrato teniendo en cuenta de este modo la variabilidad o falta de homogeneidad de la subpoblación

El tamaño de la muestra se ve afectado por diversos factores los más relevantes son los siguientes:

- La homogeneidad o heterogeneidad de la población. Cuanto más homogénea sea la población o mayor grado de semejanza tenga sobre la característica bajo estudio, menor será el número de casos necesarios para obtener una muestra que la represente con precisión y, al contrario, cuanto más heterogénea sea, más casos harán falta.
- El número de categorías, clases y el número de casos. Cuanto mayor sea el número de subgrupos, mayor tendrá que ser la muestra total necesaria para obtener su medición estadística precisa. El tamaño de la muestra que se elija ha de ser lo bastante grande para permitir evaluar con precisión las clasificaciones importantes.
- El tamaño de la muestra y la clase de muestreo. Bastará una muestra más pequeña cuando se emplea la estratificación en vez de la selección al azar ilimitada. Esto es porque la estratificación reparte el total relativamente heterogéneo en cierto número de submuestras individualmente homogéneas. Cuanta más heterogénea sea la población, mayor será la posibilidad de reducir el número de casos mediante la estratificación.
- El tamaño de la unidad de muestreo. Las grandes unidades de muestreo resulta inútiles cuando la tabulación de datos ha de referirse a familias o a individuos.

En los métodos para la selección de la muestra se busca que no invaliden la muestra para ello debe mantener el carácter representativo del universo y la validez de la inducción estadística. Los métodos de selección al azar más comunes son los siguientes:

- Selección por sorteo: Para este método se enumera correlativamente la totalidad del universo y se procede más o menos similarmente a como se realiza un sorteo de lotería preparándose bolitas u otro método similar que representan el universo y que son introducidas en un cubículo, bombo, etc., las cuales deben ser mezcladas y extraídas al azar tal y como se efectúa en sorteo cualquiera. Los números extraídos corresponden a la enumeración establecida previamente en la lista, constituyendo la muestra.
- Selección a intervalos regulares: Este sistema proporciona una muestra más representativa que el sorteo ya que garantiza la obtención de un corte transversal de la muestra. Existen diversos tipos de selección, entre los cuales se tiene el muestreo de una lista o fichero a intervalos regulares; selección por los puestos que ocupan los elementos en las páginas o en

numeración por líneas; selección por medio de un cuadrículado en el mapa, etc.

- Selección por parte de los entrevistadores: Este procedimiento tiene el riesgo de introducir propensiones o vicios al sesgo del seleccionador, por lo tanto no se recomienda.
- Uso de la tabla de números aleatorios: Es el método más recomendado ya que garantiza igual probabilidad de selección.

La fórmula que se aplica en el caso de tener una población finita sería de acuerdo a la ecuación (1)

$$n = \frac{NZ^2S^2}{(N-1)e^2 + Z^2S^2} \quad (1)$$

La selección de la muestra en una población finita (N) sin reemplazamiento en un muestreo aleatorio simple de cada estrato depende de tres parámetros:

- Uno de los parámetros es el nivel de confianza ( $Z_{\alpha/2}$ ) de que la estimación se ajuste a la realidad dentro de la hipótesis de una distribución normal, en general se establece un nivel de confianza del 95% por lo cual se asume que  $Z_{5\%/2}=1,96$ , si se quisiera tener un 100% de confianza eso implicaría un censo y no una muestra.
- El segundo parámetro es conocido como el nivel de error (e) que mide la probabilidad de aceptar una hipótesis que sea falsa como si fuera verdadera o de rechazar una verdadera como si fuera falsa. El nivel de error da una medida de con qué probabilidad una estimación basada en la muestra se aleja del valor obtenido en el censo. Para tener un nivel de 0% de error la muestra debe coincidir con la población. Los niveles de error óptimos se encuentran entre un 4 a 6% (no son complementarios con el nivel de confianza), pero se ajustan de acuerdo a la disponibilidad de recursos que disponga el investigador aceptándose como límite máximo un 40% de error.
- El tercer parámetro es la variabilidad de la muestra, esta es la proporción o probabilidad con que aceptamos un ensayo previo, normalmente p determina la proporción (%) con que se acepta la hipótesis (variabilidad positiva) y  $q=1-p$ ; representa la probabilidad (%) con que se rechaza la hipótesis (variabilidad negativa), la máxima variabilidad ocurre cuando  $p=q=0,5$ . La varianza de la muestra ( $S^2$ ) viene dada por el producto de las variabilidades  $S^2=p*q$ .

Los pasos necesarios para determinar el tamaño de una muestra empleando el muestreo aleatorio simple sí se asume que la población es conocida pero muy grande ( $N \rightarrow \infty$ ) y se quiere estimar la media. Para ello es necesario partir de dos supuestos: en primer lugar el nivel de confianza al que se quiere trabajar; en segundo lugar, cual es el error máximo que se está dispuestos a admitir en la estimación, dado estos supuestos se utiliza la ecuación (2).

$$n_{\infty} = \frac{Z_{\alpha/2}^2 S^2}{e^2} \quad (2)$$

En este caso se debe evaluar si  $N > n_{\infty}(n_{\infty} - 1)$  si esta condición se cumple el proceso termina aquí, y ese es el tamaño adecuado que se debe muestrear. Si no se

cumple, se pasa a una fase siguiente que es el obtener el tamaño de la muestra según la siguiente fórmula de la ecuación (3).

$$n = \frac{n_{\infty}}{1 + \left(\frac{n_{\infty}}{N}\right)} \quad (3)$$

Para calcular el tamaño de muestra para la estimación de proporciones poblacionales se ha de tener en cuenta los mismos factores que en el caso de la media. La fórmula de la ecuación (4) que permitirá determinar el tamaño muestral es la siguiente:

$$n = \frac{N \cdot Z_{\alpha/2}^2 \cdot P(1-P)}{(N-1)e^2 + Z_{\alpha/2}^2 \cdot P(1-P)} \quad (4)$$

Donde; **Z** correspondiente al nivel de confianza elegido, **P** es proporción de una categoría de la variable, **e** es el error máximo. **N** es el tamaño de la población

Existen otras fórmulas y criterios a aplicar pero para ello se recomienda revisar la siguiente literatura:

- Azorín, F. y Sánchez-Crespo, J.L. (1986). Métodos y Aplicaciones del Muestreo. Madrid: Alianza Editorial.
- Blalock, H. (1994). Estadística Social. México: Fondo de Cultura Económica.
- Briones, G. (1990). Métodos y Técnicas de Investigación para las Ciencias Sociales. México: Editorial Trillas.
- Gabaldón Mejía, N. (1973). "Algunos Conceptos de Muestreo". Universidad Central de Venezuela. Facultad de Ciencias Económicas y Sociales. División de Publicaciones, Caracas.
- Hoel, P. G. (1976). Introducción a la Estadísticas Matemáticas. Editorial Ariel, Barcelona.
- Raj, D. (1992). Teoría del Muestreo. México: Fondo de Cultura Económica.

### 3.3. ¿Cómo se investigará?

En todo proceso de investigación se debe tener claridad no solo del problema sino del tipo de investigación: si es exploratoria, descriptiva o explicativa. Esto definirá los métodos y su alcance, para ello se debe responder ¿qué se investiga?, ¿cuáles son las características del problema?, ¿qué resultados pueden obtenerse? y ¿cuáles son sus limitantes?. Esto debe estar claramente definido en el diseño metodológico que es parte del marco teórico y de la propuesta de investigación.

#### 3.3.1. Valor teórico

El método de análisis multivariante apunta a obtener un valor teórico que resulta de una combinación lineal de variables con ponderaciones determinadas empíricamente, para ello todas las variables deben ser aleatorias y estar relacionadas de tal forma que sus diferentes efectos no puedan ser interpretados separadamente con algún sentido. El investigador especifica, a partir del marco teórico y de las hipótesis derivadas, las variables que constituyen el modelo. El valor teórico o variable explicada (dependiente) por k variables ponderadas (de X1 a XK) puede ser expresada matemáticamente como:

$$\text{Valor Teórico} = \omega_1 X_1 + \omega_2 X_2 + \omega_3 X_3 + \dots + \omega_K X_K$$

Donde  $X_i$  (para  $i = 1$  a  $k$ ) es la variable observada que explica el valor teórico y  $\omega_i$  es la ponderación (o coeficientes) determinada por la técnica de estimación multivariante.

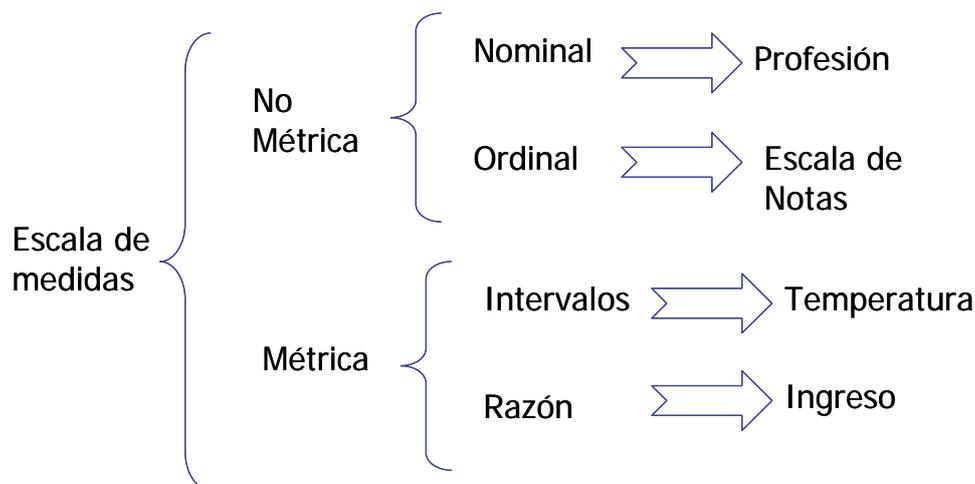
El resultado es un valor único que representa una combinación del conjunto de variables que mejor se adaptan al objeto del análisis, donde se valora el impacto conjunto y la contribución de cada variable separada al efecto del valor teórico en su conjunto.

### 3.3.2. Escala de medida

En el análisis de los datos implica separar, identificar y medir las variables del estudio que represente con precisión el concepto de nuestro interés. En la figura 7 se muestra las diferentes escalas de medidas de las variables. Las variables se miden por el tipo de dato; que pueden ser de tipo *no métrico* o cualitativo y *métrico* o cuantitativo. Los datos cualitativos o no métricos son atributos, o propiedades categóricas que describen un sujeto, por ejemplo<sup>20</sup>, si uno es hombre, no puede ser mujer. No hay cantidad de "género", sólo la condición de ser hombre o ser mujer. En este caso se tiene un dato no métrico nominal referido al género.

Las medidas *no métricas* pueden tener escalas nominales u ordinales. Una *escala nominal* asigna un número que se usa para etiquetar o identificar un sujeto u objetos como, por ejemplo, ser católico se le asigna el número 1 y ser protestante se le asigna el número 2, ser musulmán se le asigna el número 3, ser judío se le asigna el número 4 y ser no creyente se asigna el número 5. En una *escala ordinal* se busca establecer un orden, una escala categórica donde su distancia no tiene mayor relevancia sino en cuanto el establecer una jerarquía de mayor a menor importancia o preferencia, por ejemplo los niveles de confianza de un cliente frente al servicio, o la escala de notas de evaluación de un curso.

**Figura 8**



<sup>20</sup> Ob. Cit.

Fuente: Elaboración propia.

Las *escalas métricas* proporcionan un nivel mayor de precisión que las escalas no métricas ya que permiten realizar casi todas las operaciones matemáticas. Las escalas métricas son de dos tipos las escalas de intervalos y las escalas de razón, estas dos escalas tienen unidades constantes de medidas de forma que la diferencia entre dos puntos adyacentes de cualquier parte de la escala, son iguales. La diferencia entre las escalas de intervalo y las de razón es que las de intervalo tienen un punto cero arbitrario, mientras que las escalas de razón tienen un punto de cero absoluto. Las escalas de intervalos más comunes son las de temperatura como la escala Celsius y Fahrenheit. Las escalas de razón representan una medida de mayor precisión, dado que poseen todas las ventajas de las escalas inferiores más un punto de cero absoluto.

La importancia de entender los diferentes tipos de escalas de medidas se debe a dos razones, en primer lugar, el investigador debe identificar la escala de medida empleada de cada variable de forma que no se estén utilizando datos no métricos como si fueran métricos y en segundo lugar, sirve para determinar qué técnica de análisis multivariante es la más conveniente para los datos, lo cual debe ser considerado tanto para la variable dependiente como para las variables explicativas.

### 3.3.3. Error de medida, validez y fiabilidad

Un tema complementario a la escala es el error de medida que consiste en que los valores observados no son representativos de los verdaderos valores. El error de medida puede surgir por imprecisión en la medición, por ejemplo; por proponer más categorías de respuesta que las que el entrevistado promedio puede distinguir. También surge por la incapacidad de los encuestados a proporcionar información precisa, por ejemplo, el nivel preciso de renta del hogar. El impacto del error de medida es introducir “ruido” en las variables observadas, por lo cual la variable medida contendrá un valor “verdadero” más un “ruido” (Variable observada = Valor verdadero + Ruido).

Al valorar el grado de error de medida se debe considerar la validez y la fiabilidad de la medida. La validez es el grado en que la medida representa con precisión lo que se supone que representa. Para ello se debe conocer en profundidad lo que se va a medir y sólo entonces realizar la selección de la medida “correcta” de la forma más precisa posible, sin embargo la precisión no asegura la validez ya que se puede medir con precisión una variable errada y entre varias medidas posibles se debe elegir la de mayor fiabilidad.

Garantizada la validez se debe considerar la fiabilidad de las medidas. La fiabilidad es el grado en que la variable observada mide el valor “verdadero” y esta “libre de error”<sup>21</sup> y por lo tanto es lo opuesto al error de medida. Por tanto las medidas más fiables mostrarán una mayor consistencia que las medidas menos fiables

En algunos estudios se construyen *medidas compuestas* para representar un concepto, al representar un concepto a través de varias variables combinadas como indicadores se trata de representar el concepto desde varias perspectivas dando una visión más completa. El uso de indicadores múltiples permite llegar a una especificación

---

<sup>21</sup> Ob. Cit.

más precisa de las respuestas deseadas y no deja la fiabilidad a una única respuesta sino a en la respuesta “típica” de un conjunto de respuestas relacionadas.

Si bien los resultados inadecuados no dependen siempre del error de medida, la presencia del error de medida es garantía de distorsión en las relaciones observadas y reduce significativamente las técnicas de estimación ya que se aplica el principio de “Garbage In, Garbage Out”<sup>22</sup>. El reducir el error de medida puede mejorar los resultados y reforzar relaciones ya probadas.

#### 3.3.4. Orientaciones Adicionales

La primera observación es que el análisis multivariante no sustituye al análisis teórico conceptual sino que lo complementa un buen modelo conceptual orienta más que el cazar relaciones sin ningún fundamento teórico. En la figura 8 se muestra algunos esquemas donde se esquematiza las relaciones causales y sus diagramas de relaciones. Esto permite visualizar tanto el tipo de relaciones como las técnicas estadísticas a aplicar en la estimación de los coeficientes o ponderadores para obtener el valor estimado o valor teórico de la relación.

Adicionalmente, se debe guardar como criterios del estudio que las relaciones tenga significación práctica, por ejemplo, si se estima un modelo de regresión donde las variables consideradas son todas significativas estadísticamente al 95% pero su poder explicativo es muy bajo, por ejemplo, con un coeficiente de determinación del 10% , este no puede ser un modelo confiable para tomar decisiones.

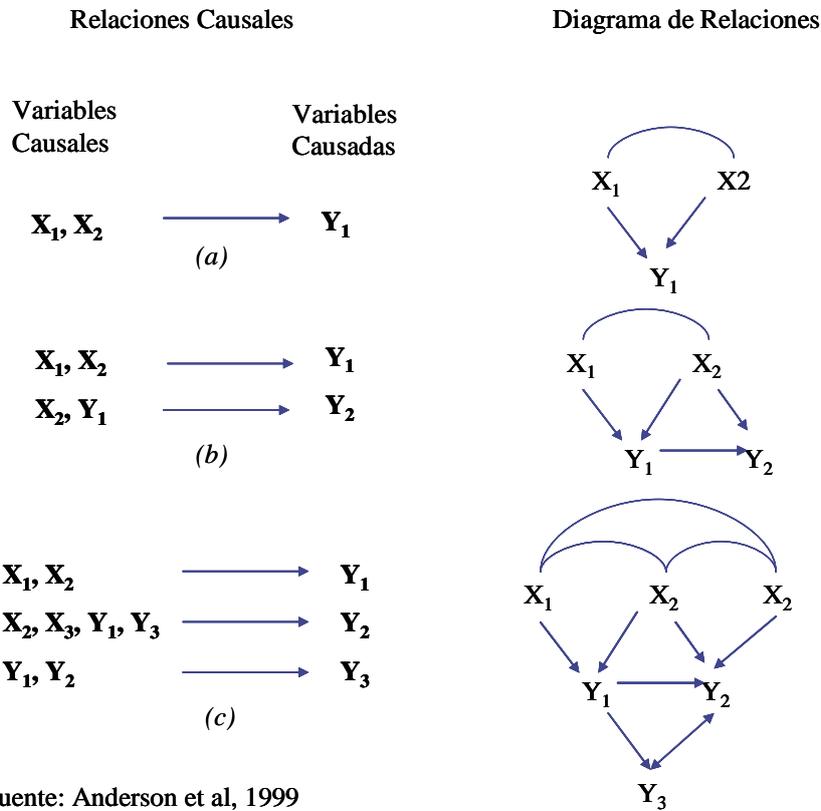
Lo anterior no sirve de razón para incluir en el modelo variables indiscriminadamente ya que introduce errores de especificación y sobre ajusta el modelo reduciendo su capacidad de predicción sólo a la muestra, sin ningún poder de predicción o replicación fuera de la muestra hacia la población bajo estudio, con la posibilidad adicional de generar un problema de multicolinealidad<sup>23</sup> en la estimación. Por ello se plantean dos principios adicionales como son la parsimonia que es poder explicar lo más con lo menos y el principio de hacerlo lo más sencillo posible que se resume en el aforismo de “hágalo estúpidamente simple o ¡estúpido hágalo simple!”

---

<sup>22</sup> Aforismo que dice que si “introduces basura en el sistema lo que obtendrás es basura”, GIGO.

<sup>23</sup> Es la creación de una dependencia lineal entre las variables explicativas que representa más dimensiones espaciales de que las que corresponde al modelo, en general la multicolinealidad es más un problema de diseño que de muestra

**Figura 9**



Fuente: Anderson et al, 1999

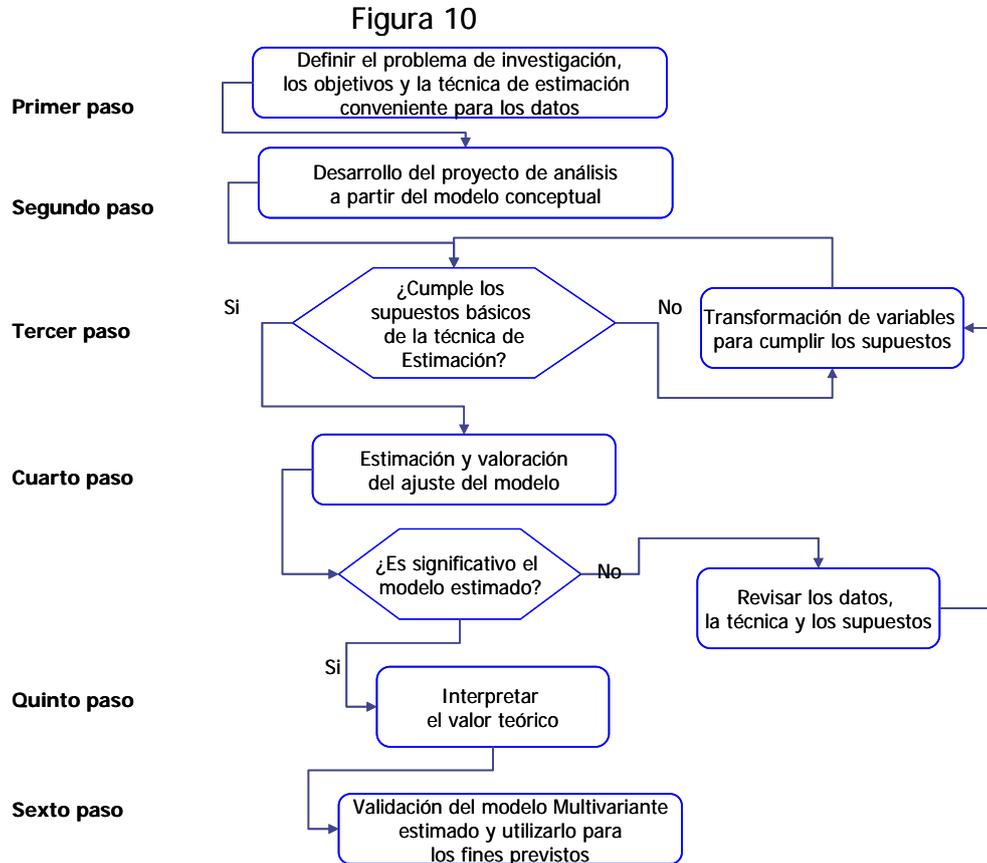
El tamaño de la muestra o el número de observaciones influye en la potencia estadísticas ya que existe una relación directa entre el tamaño de las observaciones y la potencia de las técnicas de estimación de las relaciones y hace que las pruebas de contraste sean más sensibles y hace que cualquier variable tenga un coeficiente significativo por ello los resultados deben evaluarse a la luz de la significancia practica de los datos y del propósito de la estimación.

La capacidad predictiva del modelo es la primera prueba de validez de las relaciones estimadas y eso se estudia a través de los errores de la predicción que es la distancia entre el valor observado de la muestra y el valor teórico construido a través de la relación estimada, el analista debe utilizar los errores de predicción como un punto de partida para diagnosticar la validez de los resultados obtenidos y como un indicador de las relaciones que quedan por explicar. Lo relevante es recordar que el objetivo no es obtener el mejor "ajuste" para la muestra sino desarrollar un modelo significativo que describa la población en su conjunto.

### 3.4. Etapas o pasos generales

El esquema de proceso que se propone como una aproximación general para el diseño y estimación del modelo consta de seis pasos mínimos. La aproximación se centra en el análisis de un plan de investigación bien definido comenzando con un modelo conceptual que detalle las relaciones a examinar para iniciar los trabajos empíricos, incluyendo la selección de la técnica específica y su proceso de aplicación a los datos

recopilados para el problema propuesto. Esto se resume en la figura 10. Este esquema es lo suficientemente amplio como para considerar no sólo el uso del análisis de regresión sino otras técnicas de análisis.



Fuente: Elaboración propia.

### 3.4.1. Definir el problema de investigación y la técnica multivariante conveniente

El punto de partida para cualquier análisis multivariante es definir el problema de investigación y los objetivos analíticos en términos conceptuales, antes de especificar cualquier variable o medida. Es clave el modelo conceptual o propuesta, definido en términos de: categorías de análisis, identificación de factores y relaciones, especificando los conceptos dependientes e independientes, antes que las variables. El desarrollo de un modelo conceptual no es dominio exclusivo del campo académico; se justifica, aún más, en el mundo laboral donde las decisiones tienen un mayor costo.

Con los objetivos y el modelo conceptual definido, el analista sólo tiene seleccionar la técnica multivariante adecuada. Luego de seleccionado el método

de acuerdo al tipo de relaciones de dependencia e interdependencia, se seleccionan las variables de acuerdo a los conceptos que se quieren estudiar.

La regresión múltiple es el método de análisis más usual cuando el problema consiste en establecer una relación entre una variable explicativa que se supone está relacionada con una o más variables independientes. El fin es predecir los cambios en la variable dependiente en respuesta a cambios en varias variables independientes. Este fin se logra frecuentemente a través de la técnica de los mínimos cuadrados ordinarios, con el objeto de establecer la constancia de relaciones estructurales entre variables, evaluar la validez o éxito de una política y realizar un pronóstico puntual por ejemplo, sobre el valor del dólar al cierre de un ejercicio.

Para lograr el propósito de vincular una variable dependiente (o criterio) con un conjunto de variables explicativas, se debe alcanzar dos objetivos; el primero, es maximizar la potencia conjunta de predicción de las variables independientes como una combinación lineal y en segundo lugar, comparar dos o más conjuntos de variables independientes para buscar la de mayor poder predictivo con el menor número de variables.

#### 3.4.2. Desarrollo del proyecto de análisis

Establecido el modelo conceptual, se desarrolla el estudio de acuerdo al plan de trabajo y de acuerdo a la técnica se evalúan los supuestos en que se sustenta la técnica de estimación. Los supuestos van desde la definición del tamaño mínimo de la muestra<sup>24</sup>, pasando por los tipos de variables permitidas (métricas y no métricas) y los supuestos del método de estimación<sup>25</sup> como linealidad y normalidad, entre otros aspectos.

En esta etapa se construye la base de datos y se construye las variables compuestas, se realizan análisis univariantes de los datos<sup>26</sup> y se evalúan los supuestos subyacentes del modelo y de la técnica multivariante a través de análisis de relaciones<sup>27</sup> entre las variables.

##### 3.4.2.1 Tamaño muestral y potencia estadística

La potencia estadística de la regresión múltiple se refiere a la probabilidad de detectar un nivel significativo de la relación, usualmente se utiliza como medida el coeficiente de determinación ( $R^2$ ) que es un valor que va de 0% a 100% ( $0 \leq R^2 \leq 1$ ), este es sensible al tamaño muestral y al número de variables, ya que a mayor tamaño muestral y mayor número de variables explicativas el coeficiente de determinación aumenta y tiende a uno, por este hecho se utilizan otras medidas de bondad de ajuste como; el coeficiente de determinación ajustado combinados con el valor de la función de verosimilitud ("*Log likelihood*"), el criterio de

---

<sup>24</sup> Por ejemplo, en el modelo de regresión lineal lo ideal es tener 20 observaciones por variable, en un modelo de componentes principales lo ideal es tener 50 observaciones por variables.

<sup>25</sup> Por ejemplo para el modelo de regresión, se exige linealidad en los parámetros, independencia entre las variables explicativas o no multicolinealidad, varianza constante a lo largo de la muestra u homoscedasticidad y ausencia de relación entre los errores estimados o no autocorrelación

<sup>26</sup> Estadísticos descriptivos como la media, la varianza, la asimetría y la concentración

<sup>27</sup> Matriz de covarianza y correlación, valores característicos, etc.

información Akaike (AIC), criterio de predicción del logaritmo de Amemiya (APC) y el criterio bayesiano de información de Schwarz (SBI).

El tamaño muestral no sólo afecta la potencia estadística, sino también la capacidad de extrapolar los resultados, esto está vinculado directamente con los grados de libertad del modelo, cuanto mayor sean sus grados de libertad mayor será la capacidad de generalizar los resultados, para ello se recomienda una regla heurística: de guardar una relación mínima de 5 observaciones por variable incluida. Por ejemplo, un modelo de dos variables explicativas y una variable explicada sería de 15 observaciones, el nivel deseado debería estar en una relación de 20 y el nivel óptimo es de 50 observaciones por variable, en este caso; en un modelo de 2 variables explicativas y una dependiente sería de 150 observaciones.

#### 3.4.2.2. Creación de variables adicionales

La relación básica del modelo de regresión múltiple es de una asociación lineal de datos métricos, para estimar un conjunto de coeficientes o pendientes. Si los datos no son métricos sino expresan un orden o una presencia de un evento o por el contrario la relación de las variables métricas es no lineal. En estos casos para incorporar esta variable se requiere *transformarlas*. La transformación de los datos ofrece un medio de modificar tanto las variables independientes como las variables explicativas para mejorar o modificar la relación entre las variables y permitir el uso de variables no métricas en la creación de la significancia conjunta de la relación. La transformación de los datos responde a razones “*teóricas*” (por la naturaleza de los datos) o “*empíricas*” (transformaciones producto del patrón de los datos).

Cuando los datos son del tipo categorico, por ejemplo; sobre la situación civil de la relación de pareja o el sexo del entrevistado, se incorpora a través de las variables ficticias o binarias (“*dummies*”) donde la variable *categorica de referencia* asume un valor que generalmente es uno (1) el grupo de referencia y de cero (0) el grupo omitido. Un método alternativo de codificación de la variable ficticia es la *codificación de efectos*. En este caso el grupo omitido se codifica con el valor menos uno (-1) en lugar del cero, así, los coeficientes representan diferencias para cualquier grupo respecto de la media de todos los grupos y no la media respecto al grupo omitido. La diferencia entre uno y otro modelo se encuentra en la interpretación de los coeficientes de la variable ficticia.

Otra transformación usual es la de convertir una relación curvilínea en una relación lineal para ello se utilizan dos tipos de transformaciones las logarítmicas y de polinomios. Estas últimas, son muy útiles en transformaciones en relaciones curvilíneas simples y en especial en las de tipo de relaciones univariantes, existen métodos de comparar formas funcionales pero no se garantiza que el elegido sea la adecuada a la población, aunque lo sea para la muestra. Por ejemplo una relación lineal es un polinomio de grado uno y un polinomio de grado dos es una función cuadrática; por ejemplo:

$$Y = b_0 + b_1 X_1 + b_2 X_1^2$$

Donde  $b_i$  son los coeficientes (para;  $i = 0,1,2$ ) y  $X_i$  son las variables. El término cúbico o potencia de tercer grado es la mayor potencia utilizada habitualmente. Los criterios de Amemiya, Akaike y Schwarz, sirven para elegir el

grado de polinomio adecuado. El potencial problema de los polinomios es la de introducir multicolinealidad en los estimadores y reducir los grados de libertad.

Cuando una variable independiente se ve afectada por otra variable independiente cambiando su forma se dice que existe un efecto moderador, esto se expresa como un *efecto interacción* y es similar al término de interacción que se encuentra en el análisis multivariante de la varianza. En el análisis de regresión el término moderador es una variable compuesta formada por la multiplicación de  $X_1$  por el moderador  $X_2$ , que entra en la ecuación de regresión. La relación con efectos de interacción se “modera” como:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_1 X_2$$

Donde:  $b_0$  es la constante;  $b_1 X_1$  es el efecto lineal de la primera variable,  $b_2 X_2$  es el efecto lineal de la segunda variable y  $b_3 X_1 X_2$  es el efecto moderador de la primera sobre la segunda. Para comprobar si el efecto moderador es significativo, se debe estimar primero la ecuación sin moderar y luego estimar la ecuación con interacción. Si el cambio en el poder explicativo es estadísticamente significativo, entonces el efecto moderador debe ser incorporado en el modelo.

### 3.4.3. Evaluación de los supuestos básicos de la técnica de estimación

Las técnicas basadas en la inferencia estadística deben tener en cuenta los supuestos de normalidad multivariante, como la linealidad, independencia de los términos de error e igualdad de las varianzas en una relación de dependencia. Antes de realizar cualquier pronóstico del modelo se debe asegurar el cumplimiento de los supuestos estadísticos y conceptuales.

Estos análisis deben realizarse también después de estimado el modelo de regresión en el paso siguiente. La contrastación de los supuestos deben realizarse no sólo en la fases iniciales sino también después de que el modelo sea estimado. El criterio es sí en el proceso de cálculo de los coeficientes y predicción de la variable dependiente, se cumplen los supuestos del análisis de regresión. Los supuestos que se examina son los siguientes:

- La linealidad del fenómeno medido
- La varianza constante del termino de error
- La independencia de los términos de error
- La normalidad de la distribución del termino error

La principal medida del error de predicción es el residuo que es la diferencia entre el valor observado y el valor estimado como variable criterio. El examen de los gráficos de residuos y de las variables independientes y de las predicciones son básicos en la identificación de los incumplimientos de los supuestos para el conjunto de la relación. Usualmente se utilizan los residuos estandarizados con el fin de hacerlos comparables entre modelos y especificaciones. El grafico más habitual se construye con los residuos en el eje de las ordenadas y los valores predichos en el eje de las abscisas Existen adicionalmente un conjunto de pruebas formales que son propias para evaluar cada supuesto.

### 3.4.4. Estimación del modelo y valoración del ajuste del modelo

Satisfecho los supuestos del modelo, se estima efectivamente el mismo ya sea minimizando los errores, o la distancia, o maximizando la probabilidad de ocurrencia del evento. Luego se realiza la valoración global del ajuste del modelo, el orden de evaluación debería seguir el siguiente orden.

- i. Evaluación de los supuestos teórico de las relaciones esperadas en la estimación de los coeficientes, en este sentido se analiza el signo de la relación y el peso relativo del mismo en el modelo
- ii. Evaluación de la significancia estadística de los coeficientes estimados y de las relaciones globales o de interdependencia del modelo
- iii. Evaluación de los supuestos de la técnica multivariante empleada

La búsqueda del mejor modelo estará dado por el que describa mejor la población en su conjunto y su por su capacidad predictiva en cuanto a replicar valores fuera de la muestra.

Existen diversos procesos de búsqueda secuencial los más usuales son:

- i. La estimación por etapas o paso a paso ("*stepwise*")
- ii. La adición progresiva ("*forward*")
- iii. La eliminación progresiva. ("*backward*")

Por esto métodos se analizan la robustez y estabilidad de los parámetros estimados y su sensibilidad al sacar o introducir variables u observaciones. Si existe una alta sensibilidad a ciertas observaciones se deben analizar esas observaciones influyentes.

La estimación paso a paso o por etapas se realiza de la siguiente manera:

- Se construye una matriz de correlación parcial simple y se selecciona la variable que tenga el mayor coeficiente de correlación simple y se estima una regresión lineal simple con esta variable.
- Se incorpora la segunda variable de mayor correlación simple y se compara si esta incorporación contribuye significativamente a la explicación del modelo a través de una prueba F, en caso de no ser así, se saca del modelo y se incorpora la otra variable que le sigue en significación y se realiza el mismo proceso y se estima el modelo con esta nueva variable
- Esta posibilidad de eliminar e incorporar variables en forma sucesiva en el modelo es lo que la define como "*stepwise*". Con este método el investigador debe asegurar un nivel mínimo de error y trabajar con criterios conservadores de Error Tipo I del 1% al 5%

La estimación progresiva adicionando es similar a la estimación paso a paso sólo que no se eliminan variables ya incorporadas significativas. La eliminación progresiva parte del modelo más general incorporando todas las variables y luego se van eliminando las variables no significativas desde las menos significativas y se va evaluando la estabilidad del modelo y el cumplimiento de los supuestos. Una vez eliminada o incorporada una nueva variable no se debe revertir la acción posteriormente. También es posible combinar métodos de búsqueda dependiendo de la experticia o conocimiento del investigador. El principal problema potencial es de generar multicolinealidad en el proceso de búsqueda.

El modelo se selecciona de acuerdo al cumplimiento de los supuestos establecidos desde el punto de vista: Teórico o funcional, estadísticos inferenciales y del método de estimación multivariante o econométricos.

- Los supuestos teóricos, se vinculan respecto al signo esperado de la relación y a la magnitud esperada de la estimación. Por ejemplo si se

estima una función de producción se espera que el coeficiente sea positivo y exprese una relación de rendimientos decrecientes entre el insumo y el producto.

- La significación estadística de los coeficientes se evalúa a través de pruebas individuales de contraste de valores medios a través del error estándar y el coeficiente estimado utilizando la distribución “*t Student*” para muestras pequeñas o normal estandarizada, para muestras grandes.
- La significación conjunta se realiza a través de una prueba de varianza explicada utilizando el estadístico F , donde se contrasta la parte explicada versus la parte no explicada del modelo y si es mayor o no al valor en tabla para los grados de libertad respectivos y nivel de Error Tipo I establecido.

Un proceso importante es la identificación de las observaciones influyentes que pueden afectar a la  $y$  y que pueden tener su origen en a) un error en la entrada de los datos; b) una observación válida aunque excepcional que es explicable por una situación extraordinaria; c) una observación excepcional sin explicación plausible y ; d) una observación ordinaria en sus características individuales pero excepcional en su combinación conjunta. Cada caso tiene una vía de acción pero sin duda el primer caso debe ser corregido o eliminado y los otros debe ser evaluado caso a caso.

#### 3.4.5. Interpretar el valor teórico

Realizado un ajuste aceptable del modelo, toca interpretar los coeficientes individuales y el valor teórico de la relación. La interpretación se realiza examinando los coeficientes estimados tanto en su signo como su magnitud. La interpretación puede conducir a la re-especificación y re-estimación del modelo en términos de variables, formas funcionales y método de estimación. El objetivo es lograr la especificación con la mayor capacidad explicativa que puedan extrapolarse al total de la población.

#### 3.4.6. Validar el modelo estimado

Antes de convalidar los resultados se deben someter a un conjunto final de diagnóstico asegurando la capacidad explicativa y el grado de generalidad a la población de donde se extrajo la muestra. Esto se puede realizar dividiendo la muestra en submuestras, dejando una submuestra que no es incluida en la estimación y dejándola, a esta submuestra para pronosticarla con el modelo estimado y con ello ver si esta submuestra no incluida es predecible o no con coherencia utilizando el modelo estimado. Además se espera que el modelo estimado sea inclusivo de otros modelos explicativos, es decir que tenga mayor capacidad de incluir varias teorías competitivas y poder contrastarla una versus otras.

## Anexo 1

### Ventajas y Limitaciones del Uso del Muestreo

<b>Ventajas</b>	<b>Limitaciones</b>
a) Economía de tiempo y dinero en todas las etapas de una operación estadística	a) No permite hacer cálculos, proyecciones o tabulaciones con respecto a áreas, grupos o sectores pequeños de una población cuando se requiere información más precisa
b) Ampliación del alcance de una investigación	b) Se puede generar una menor precisión sobre las características de la población y su capacidad predictiva
c) Mejor entrenamiento, formación y control del personal que interviene en la investigación	c) El cálculo de los resultados está sujeto a los errores del muestreo en sí, además de los errores que puedan contener los datos y que inclusive aparecen en una enumeración total (errores ajenos al muestreo)
d) Mayor control de la integridad y calidad de la enumeración	d) Existe la creencia que las investigaciones completas ofrecen un margen más sólido y eficiente de exactitud, que los resultados, basados en la muestra
e) Mayor rapidez y eficiencia en el procesamiento y publicación de los datos	e) En base a una muestra se dificulta la tabulación cruzada de algunas características investigadas
f) Menor demanda de personal y de otros recursos	f) La probabilidad del error se multiplica en las diferentes operaciones que surjan de su propio mecanismo
g) Verificaciones posteriores a una investigación	g) La preparación estadístico matemática del muestrista (que selecciona la muestra y los instrumentos) para el diseño eficiente de la muestra en todas sus etapas
h) Mayor confiabilidad y precisión en los datos obtenidos	

Fuente: Gabaldon Mejia, Nestor. 1973. "Algunos Conceptos de muestreo" Universidad Central de Venezuela. Facultad de Ciencias Sociales. División de publicaciones. Caracas, Venezuela.

## **REFERENCIAS Y BIBLIOGRAFÍA**

- Anderson, Black, Tathan y Hair (1999) **Análisis Multivariante**. Prentice Hall, Madrid
- Azorín, F. y Sánchez-Crespo, J.L. (1986). **Métodos y Aplicaciones del Muestreo**. Madrid: Alianza Editorial.
- Blalock, H. (1994). **Estadística Social**. México: Fondo de Cultura Económica.
- Briones, G. (1990). **Métodos y Técnicas de Investigación para las Ciencias Sociales**. México: Editorial Trillas.
- Ferrater Mora, José (1980). **Diccionario de Filosofía Abreviado**. Editorial Sudamericana, Buenos Aires
- Gabaldón Mejía, N. (1973). “**Algunos Conceptos de Muestreo**”. Universidad Central de Venezuela. Facultad de Ciencias Económicas y Sociales. División de Publicaciones, Caracas.
- Gilson, Étienne (197). **El Realismo Metódico**. Ediciones Encuentro, Madrid
- Giussani, Luigi (1987). **El Sentido Religioso**. Ediciones Encuentro, Madrid
- Hoel, P. G. (1976). **Introducción a la Estadísticas Matemáticas**. Editorial Ariel, Barcelona.
- Pineda, E. B., E. L. de Alvarado y †F. H. de Canales (1994). **Metodología de la Investigación**. Organización Panamericana de Salud. Washington, D. C.
- Raj, D. (1992). **Teoría del Muestreo**. México: Fondo de Cultura Económica.
- Real Academia Española, (1992). **Diccionario de la Lengua Española, vigésima primera edición**. Editorial Espasa Calpe, S. A. Madrid.
- Seco, Manuel, (1993a). **Gramática esencial del español**. Introducción al estudio de la lengua. Espasa Calpe, Madrid.
- Seco, Manuel, (1993b). **Diccionario de dudas y dificultades de la lengua española**. Espasa Calpe, Madrid.