



sincie

Sociedad Hispana de
Investigadores Científicos



JOSÉ SUPO

**CÓMO SE HACE
UNA TESIS**

2DA EDICIÓN. LIMA. BIOESTADISTICO. 2026.

Cómo se hace una tesis

Para tesis de pregrado, maestría y doctorado

JOSÉ SUPO

Sociedad Hispana de Investigadores Científicos

BIOESTADISTICO

2026

CRÉDITOS

Título de la publicación:

Cómo se hace una tesis: Para tesis de pregrado, maestría y doctorado

Número de edición:

Segunda edición

Autor:

José Antonio Supo Condori

Editorial / Sello:

BIOESTADISTICO

Razón social y domicilio legal del editor e impresor:

BIOESTADISTICO Empresa Educativa EIRL, Av. Sánchez Carrión 615,
Oficina 1008, Jesús María, Lima, Perú.

Mes y año de edición:

Enero de 2026

Tiraje de la publicación (edición Perú):

1000 ejemplares

Depósito Legal:

Biblioteca Nacional del Perú N° 2025-14954

ISBN:

978-612-99163-1-6

Derechos y licencia

© BIOESTADISTICO, 2026. Esta obra está bajo una licencia Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional (CC BY-NC-ND 4.0).

Usted es libre de compartir este libro, copiarlo y redistribuirlo en cualquier medio o formato, siempre que: atribuya correctamente la autoría, no haga uso comercial y no genere obras derivadas.

Más información en: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Cómo citar este libro

APA 7ª edición:

Supo, J. A. (2026). *Cómo se hace una tesis: Para tesis de pregrado, maestría y doctorado* (2.ª ed.). BIOESTADISTICO.

Vancouver:

Supo JA. *Cómo se hace una tesis: Para tesis de pregrado, maestría y doctorado*. 2ª ed. Lima: BIOESTADISTICO; 2026.

Nota de impresión

Edición Perú (impresión local): Impreso en el Perú

Edición internacional (Amazon KDP): Printed in the USA by Amazon KDP

CONTENIDO

1	Capítulo 1. La investigación	Pág. 1
2	Capítulo 2. El estudio	Pág. 8
3	Capítulo 3. El marco teórico	Pág. 17
4	Capítulo 4. Las variables	Pág. 24
5	Capítulo 5. Los objetivos	Pág. 33
6	Capítulo 6. La hipótesis	Pág. 40
7	Capítulo 7. El muestreo	Pág. 47
8	Capítulo 8. La recolección de datos	Pág. 56
9	Capítulo 9. Los instrumentos	Pág. 65
10	Capítulo 10. Los resultados	Pág. 74

CAPÍTULO 1: LA INVESTIGACIÓN

La investigación es una actividad orientada a producir conocimiento fundamentado sobre la realidad, mediante estudios intencionalmente planteados y metodológicamente estructurados, cuyos resultados pueden expresarse en distintos productos formales, como la tesis, el artículo científico o el reporte técnico, según el propósito y el contexto de aplicación.

Los tipos de investigación se clasifican a partir de criterios metodológicos fundamentales: la intervención del investigador, el control de la medición, el número de ocasiones en que se realiza la medición y el número de variables analíticas. Esta clasificación permite diferenciar estudios observacionales y experimentales, prospectivos y retrospectivos, transversales y longitudinales, así como descriptivos y analíticos.

Los niveles de investigación constituyen una progresión cognoscitiva que comprende los niveles exploratorio, descriptivo, relacional, explicativo, predictivo y aplicativo. Cada uno responde a una intención específica, que va desde la identificación y comprensión inicial de los fenómenos hasta la intervención orientada a mejorar una condición, organizando de manera lógica y secuencial el desarrollo de la línea de investigación.

Los diseños de investigación traducen el propósito del estudio en estrategias metodológicas concretas, entre las que se incluyen los diseños epidemiológicos, experimentales, comunitarios y de validación de instrumentos. Cada diseño implica supuestos, controles y procedimientos propios, cuya adecuada aplicación garantiza la coherencia analítica y la validez de los resultados.

Tipos de investigación

- Según la intervención del investigador: con intervención contienen a los **experimentales** y sin intervención son denominados **observacionales**.
- Según el control de la medición de la variable de estudio: con control de la medición **prospectivos** y sin control de la medición **retrospectivos**.
- Según el número de mediciones sobre la variable de estudio: con una sola medición **transversales** y con más de una medición **longitudinales**.
- Según el número de variables analíticas: con una variable analítica **descriptivos** y con más de una variable analítica **analíticos**.

Estos criterios de clasificación deben cumplir las condiciones de: exhaustivo y excluyente

- **Exhaustivo** significa que no puede haber un estudio que no pueda ser asignado en alguno de los dos grupos.
- **Excluyente** significa que no puede haber un estudio que pertenezca a ambos grupos al mismo tiempo.

Los principios de la taxonomía: exhaustive, excluyente y parsimonia, que indica que mientras menos es mejor, el número mínimo de grupos será dos.

1. Según la intervención del investigador

El investigador puede modificar a las variables de estudio (con intervención) o simplemente observar los hechos, tal cual ocurren en la naturaleza.

Los estudios con intervención

- Son prospectivos siempre, porque planifican la medición de las variables en las unidades de estudio.
- Son longitudinales generalmente, para observar el cambio producido por la intervención.
- Son analíticos siempre, porque hay que verificar el efecto de la intervención del investigador.
- Son estudios explicativos, por lo menos (en los niveles de investigación).

Los estudios sin intervención llamados “observacionales”:

- Son los estudios exploratorios o cualitativos.
- Todos los estudios descriptivos.
- Todos los estudios relacionales.
- Pueden llegar a ser estudios explicativos siempre que sean observacionales.

2. Según el control de la medición de la variable de estudio

Con datos planificados y control de los sesgos de medición: prospectivos. Con datos que provienen de mediciones donde el investigador no participó datos secundarios: retrospectivos.

Los estudios prospectivos. Utilizan datos primarios, datos que provienen de mediciones planificadas por el investigador, por lo tanto estos datos, poseen el control de los sesgos de medición.

Los estudios retrospectivos. Utilizan datos secundarios, datos que provienen de mediciones, donde el investigador no participó. El investigador no puede avalar la precisión y exactitud de los datos.

3. Según el número de mediciones sobre la variable de estudio

Con una sola medición transversales, con más de una medición: 2, 3, 4, 5 o las necesarias: longitudinales. Son llamados también estudios de seguimiento.

Los estudios transversales. Miden a las variables, a todas en general, en una sola ocasión, bajo esta condición si realizamos comparaciones, se trata de una comparación entre grupos.

Los estudios longitudinales. Miden a la variable de estudio en más de una ocasión: lo mínimo es dos. Se trata de la comparación antes-después, el análisis estadístico se denomina: de medidas repetidas.

4. Según el número de variables analíticas

Desde el punto de vista estadístico, con una sola variable: descriptivo, y si relacionamos variables: analítico.

Los estudios descriptivos. Poseen análisis estadístico univariado, pudiendo no ser inferenciales si solo describen características, e inferenciales si estiman parámetros de la población de estudio.

Los estudios analíticos. Relacionan variables, variables que aparecen en el enunciado del estudio. El análisis estadístico es bivariado si son dos variables, o multivariado si son más de dos variables.

Niveles de investigación

Una línea de investigación es una sucesión continua de estudios, que comienza con el descubrimiento de un problema y termina con el planteamiento de su solución. Para ello, la línea de investigación debe atravesar niveles:

1. Nivel exploratorio

Se plantea cuando se observa un fenómeno hecho o acontecimiento que debe ser analizado y puede perfectamente nacer de una serendipia. El estudio exploratorio: identifica los hechos, los interpreta, establece reglas para reconocer su presencia y finalmente los define operacionalmente.

2. Nivel descriptivo

Se encarga de la descripción y cuantificación del fenómeno, hecho o acontecimiento, considerando siempre una circunstancia temporal y geográfica determinada. Sus resultados pueden ser solamente descriptivos o se pueden extrapolar hacia un conjunto mayor denominado población.

3. Nivel relacional

Son estudios que plantean relación entre variables, sin importar su naturaleza. No pretenden demostrar relaciones de causalidad, sino solamente el descubrimiento de una dependencia probabilística, las relaciones encontradas aquí son útiles para plantear una hipótesis de causalidad.

4. Nivel explicativo

Son estudios que plantean relaciones de causalidad entre variables, aquí la estadística es insuficiente para completar sus objetivos, por lo que, aparece el concepto de criterios de causalidad, donde el principal criterio es la relación temporal e implica demostrar que la causa estuvo antes que el efecto.

5. Nivel predictivo

Calcula la probabilidad de ocurrencia de un fenómeno hecho o acontecimiento, parte de la premisa que ya se conocen a las variables predictoras identificadas en el nivel anterior, utiliza modelos y aprendizaje automático para predecir eventos futuros.

6. Nivel aplicativo.

Realiza una intervención sobre la población de estudio, para conseguir una mejora en alguna condición, analiza el éxito de los procesos resultados e impacto de la intervención. El análisis estadístico corresponde a las metodologías de mejora continua como el control estadístico de procesos.

Los niveles de investigación, si se pudieran describir con una sola palabra:

- Para el exploratorio: cualitativo
- Para el descriptivo: univariado
- Para el relacional: bivariado
- Para el explicativo: causalidad
- Para el predictivo: predecir
- Para el aplicativo: mejorar

Si consideras únicamente dos fases: el nivel exploratorio es el cualitativo, y todos los demás son cuantitativos.

Asimismo, el nivel exploratorio y descriptivo buscan únicamente concluir sobre el grupo estudiado, o en el mejor de los casos, sobre la población de la cual se extrajo una muestra, a esto se le denomina validez interna, mientras que los estudios analíticos buscan validez externa.

Las hipótesis, cuando no hay que demostrar relaciones de causalidad son empíricas, pero si hay relaciones de causalidad entre variables las hipótesis son racionales.

Los datos que se obtienen para realizar un estudio se denominan estáticos (matrices o bases de datos) pero, si se producen por el trabajo cotidiano o profesional se denominan dinámicos (datasets).

La investigación que busca: saber por conocer, se denomina pura o básica, pero si se pretende mejorar a la población de estudio en alguna condición se denomina aplicada.

Asimismo, los estudios pueden ser inferenciales y no inferenciales.

Dependiendo de, si queremos llevarlas las conclusiones de un grupo estudiado hacia un conjunto mayor denominado población. Todos los estudios exploratorios son no inferenciales, además del objetivo describir correspondiente al nivel descriptivo, de allí en adelante todos los estudios son inferenciales.

Finalmente, los estudios son: con y sin intervención. Si nos ubicamos en el nivel explicativo: los estudios explicativos sin intervención son denominados observacionales, pero si hay una manipulación sobre la variable de estudio se denominarán experimentales.

Diseños de investigación

En investigación, un diseño es una estrategia metodológica y estadística, para desarrollar una idea de investigación, que es lo mismo que completar un propósito investigativo. De modo que, habrá tantos diseños como ideas de investigación se nos ocurran, y sin embargo los podemos agrupar, según el campo del conocimiento donde se originaron. Siendo así:

- Los diseños epidemiológicos se originaron en las ciencias de la salud
- Los diseños experimentales en las ciencias naturales
- Los diseños comunitarios o ecológicos en las ciencias sociales y
- La validación de instrumentos en las ciencias del comportamiento.

1. Diseños epidemiológicos

Nacen en las ciencias de la salud con la intención de estudiar la morbimortalidad en seres humanos.

Los diseños epidemiológicos, pueden ser descriptivos o analíticos.

Los descriptivos incluyen a: la serie de casos, al diseño de prevalencia y al diseño de incidencia.

Los diseños analíticos, pueden ser observacionales o experimentales. Los observacionales incluyen al diseño de casos y controles y al diseño de cohortes. Dentro de los experimentales tenemos al ensayo clínico y también al ensayo de población.

2. Diseños experimentales

Nacen en las ciencias naturales y siempre tienen: manipulación y control. La manipulación se refiere a la intervención por parte del investigador a propósito del estudio, y el control: al control metodológico y estadístico.

Los experimentos, tienen como mínimo un grupo de estudio y un grupo control, elegidos de manera aleatoria de un conjunto mayor, una medida inicial o basal y una medida final, siendo así se le denomina experimento verdadero, porque cuenta con un control interno a través de las medidas repetidas, y un control externo o paralelo a partir de su grupo control.

Es posible que en algunos casos no se pueda aleatorizar, en cuyo caso se denomina cuasi experimento. Menos que eso sería contar únicamente con una medición después de la intervención en cuyo caso, se va a requerir necesariamente de un grupo control llamándose pre-experimento.

3. Diseños comunitarios

Nacen en las ciencias sociales y las variables que estudian son: agrupadas, agregadas o ecológicas, como el índice de analfabetismo o la tasa de mortalidad materna y habitualmente utilizan datos secundarios correspondientes a estudios retrospectivos.

Los diseños comunitarios o ecológicos, pueden ser: exploratorios o analíticos. Los exploratorios incluyen a los estudios etnográficos y a los grupos focales. Los analíticos pueden ser, con y sin intervención.

Con intervención incluyen a la vigilancia epidemiológica y la política pública sería otro ejemplo. Los estudios sin intervención incluyen al censo.

4. Validación de instrumentos

Nacen en las ciencias del comportamiento y tienen como primera intención la definición operacional de un constructo, de un concepto, de una propiedad subyacente que se supone posee el sujeto y su segunda intención es la medición del constructo.

Siendo así hay que crear un instrumento, redactando los ítems preguntas o reactivos, para luego validar el instrumento, evaluando sus propiedades métricas y cuando sea necesario adaptarlo, a través de una traducción lingüística, actualización o una adaptación cultural.

La creación del instrumento, asegura la validez de contenido a través de la validez racional, la validación por jueces o la validez de respuesta.

Construido así, el instrumento procedemos a la evaluación de sus propiedades métricas a través de: la validez de constructo, la fiabilidad o confiabilidad, la estabilidad a través de la repetibilidad y reproducibilidad, la validez de criterio y el rendimiento diagnóstico.

En los casos en que sea necesaria la **adaptación del instrumento**, por haberlo traducido, por haberlo actualizado, o por haber realizado una adaptación cultural.

Estos diseños no son excluyentes, un diseño epidemiológico puede ser experimental, puede ser comunitario y también puede incluir a la validación de instrumentos.

CAPÍTULO 2: EL ESTUDIO

La línea de investigación representa el eje temático que articula una sucesión continua de estudios jerarquizados. Surge de un fenómeno, condición o situación de interés, y orienta el desarrollo continuo del conocimiento. Constituye el marco dentro del cual cada estudio adquiere sentido y contribuye a un avance del conocimiento acumulativo.

El propósito del estudio expresa la intención del investigador y constituye un punto específico dentro de la línea de investigación. A partir de él se identifica el nivel investigativo, se orienta la elección del diseño y se anticipa la técnica estadística o estrategia analítica necesaria para responder la pregunta planteada.

La población de estudio delimita el conjunto de unidades sobre las cuales se obtendrá información para responder al propósito planteado. Incluye definiciones conceptuales, espaciales y temporales, así como criterios de inclusión y exclusión que aseguran validez interna. Una población correctamente delimitada determina el equilibrio entre la factibilidad y la precisión del estudio.

El enunciado del estudio reúne en una sola oración el propósito, la línea de investigación y la población de estudio, expresando de forma clara qué se va a investigar. Señala el punto donde la intención del investigador se conecta con la realidad que será estudiada y sirve como base para todo el desarrollo metodológico del trabajo.

La línea de investigación

En matemáticas una línea es una sucesión continua de puntos, en investigación cada uno de esos puntos representa a un estudio.

Una línea de investigación comienza con el descubrimiento del problema y termina con el planteamiento de su solución, para ello, hay que construir el conocimiento a través del desarrollo de un conjunto de estudios jerarquizados y ordenados; es a lo que denominamos línea de investigación.

La diabetes, la hipertensión y el cáncer de pulmón, son ejemplos de: **líneas de investigación** dentro de la medicina.

La medicina, la nutrición y la biología son ejemplos de: **áreas del conocimiento** dentro de las ciencias de la salud.

Las ciencias de la salud, las ciencias sociales y las ingenierías son ejemplos de: **campos del conocimiento**.

De manera que, una línea de investigación es un tema muy puntual dentro de un área y un campo del conocimiento.

Las líneas de investigación surgen a partir de la identificación de condiciones o situaciones:

Son condiciones problemáticas: la diabetes, el fracaso escolar y el estrés laboral.

Son situaciones no problemáticas, el rendimiento académico la calidad de vida y el clima laboral.

Las líneas de investigación estudian hechos fenómenos o acontecimientos de interés para el investigador

Existen líneas de investigación personales e institucionales:

- **Las líneas de investigación personales** son aquellas elegidas por el investigador a partir de su interés personal y su finalidad cognoscitiva o deseo de conocer, además de su relación con el tema.
- **Son líneas de investigación institucionales** aquellas desarrolladas por personas jurídicas, a partir de su responsabilidad social-empresarial. Es llevada a cabo por sus trabajadores, se trata claramente de servicios profesionales.

Para encontrar tu línea de investigación...

Haz un listado de todos los temas que te gustan dentro de tu profesión y sobre ellos, haz una selección de los que más dominas, para luego encontrar lo que necesitan tus pacientes, usuarios, clientes, trabajadores o alumnos; es decir, tu población de estudio y finalmente selecciona aquello por lo que te pueden pagar.

A partir de estos cuatro elementos, en su conjunción habrás encontrado:

Tu línea de investigación y sin embargo...

Para conocer tu línea de investigación no existe un algoritmo...

Considera las siguientes recomendaciones:

1. **Número uno.** Una línea de investigación es un talento natural, recuerda que “a la hierba no le cuesta trabajo crecer”.
2. **Número dos.** Es una vocación dentro de tu profesión “que la inspiración te encuentre trabajando” como cuando Alexander Fleming descubrió la penicilina.
3. **Número tres.** Explora las experiencias pasadas y “recuerda que nada sucede por casualidad” uno los puntos como lo hizo Steve Jobs.
4. **Número cuatro.** Es lo que tu población espera de ti porque “el que no vive para servir no sirve para vivir” la población de estudio es la razón de ser del investigador.
5. **Número cinco.** Piensa en tus expectativas personales y en la necesidad que tienes para crecer, piensa en lo que podrías llegar a ser, con la ayuda de tu línea de investigación.
6. **Número seis.** Piensa en el futuro de la línea de investigación, atrévete a cruzar las fronteras del conocimiento y explorar la estrategia del “océano azul”.
7. **Número siete.** Una línea de investigación no se elige, se asume. El agua de la montaña, tarde o temprano siempre llega al mar.

- Asúmeme en la búsqueda de tu pasión -

El propósito del estudio

Las ideas que generan intenciones, se expresan a través de: el lenguaje.

En investigación el propósito del estudio, es la expresión gramatical de la idea de investigación.

Todo trabajo de investigación tiene una intención que se expresa en el propósito del estudio, el cual genera la necesidad de, desarrollar un estudio dentro de una línea de investigación.

Imagina una línea vertical, la denominaremos **línea de investigación...**

Ahora imagina una línea horizontal, la llamaremos **nivel investigativo...**

El punto exacto donde se cortan estas dos líneas corresponde al **propósito del estudio**.

Algunos autores se refieren a él como: especificidad del estudio, otros lo conocen como finalidad cognoscitiva, también recibe el nombre de intención analítica e incluso de objetivo del estudio.

- **Especificidad del estudio**, porque se refiere al hecho específico que se quiere estudiar. Sin embargo, en validación de pruebas diagnósticas la especificidad, es la capacidad de un test de dar negativo en una persona sana, por lo que, puede causar confusiones, con el propósito del estudio.
- **Finalidad cognoscitiva**, porque su fin primario y único es conocer, y sin embargo, este concepto, no alcanza a la investigación aplicada.
- **Intención analítica**, si bien todo trabajo de investigación analiza datos, en estadística, el estudio analítico hace referencia a la relación entre variables, por lo que, para evitar confusiones, es mejor usar el nombre de: propósito del estudio.
- **Objetivo del estudio**, los objetivos son acciones que se pueden completar, mientras que el propósito del estudio es un sustantivo que expresa la idea de investigación.

El propósito del estudio es un punto concreto dentro de la línea de investigación.

Son ejemplos de propósitos investigativos: el diagnóstico, la prevalencia, los factores de riesgo, las causas, el pronóstico y el tratamiento; siempre respecto de una línea de investigación.

- El propósito **determina el nivel investigativo**, en el ejemplo: influencia del clima organizacional sobre la calidad de la atención percibida por los usuarios, el término influencia, hace alusión a una relación causa efecto, nos encontramos entonces en el nivel explicativo.
- Del propósito se **deriva el objetivo del estudio**, ejemplo de propósito: prevalencia, ejemplo de objetivo: estimar la prevalencia, claramente los objetivos son acciones que se pueden completar. Por lo tanto, corresponden a la versión operativa del propósito del estudio, que refleja la intención del investigador.
- El propósito del estudio **sugiere el diseño del estudio**, en el ejemplo: incidencia de enfermedades diarreicas en los meses de verano, el diseño de incidencia se refiere a: el estudio observacional, prospectivo, longitudinal y descriptivo. Incidencia es el propósito y también el diseño.
- El propósito **sugiere la prueba estadística** en el ejemplo: correlación entre los niveles de hemoglobina de la madre y el peso de su recién nacido, la correlación corresponde al propósito del estudio y también al procedimiento estadístico, denominado correlación de Pearson.
- El propósito **puede representar la técnica estadística** en el ejemplo: análisis de la supervivencia de los pacientes con cáncer de estómago sometidos a dos tratamientos. Análisis de supervivencia, es la técnica estadística y también el propósito del estudio.
- **El propósito del estudio por sí mismo**, según el nivel investigativo, son ejemplos de propósitos investigativos: la interpretación, la prevalencia, los factores de riesgo, la influencia, la prevención y el control; en cada uno de los niveles investigativos, según su intención.

Si al propósito del estudio, le sumamos la línea de investigación, entonces tenemos: **el tema de investigación** que, en el ámbito académico se conoce como **el tema de tesis**.

La población objeto de estudio

Una población de estudio, es un conjunto de unidades de estudio que, son objeto de interés por parte del investigador.

En todo proceso de observación, encontramos los siguientes elementos: el observador o investigador, el ente observado o población objeto de estudio, los medios de observación y las circunstancias de la observación.

El ente observado o población objeto de estudio está habitualmente conformado por un individuo o un conjunto de individuos, los llamaremos sujetos. Pero también, la población objeto de estudio, puede estar conformada por objetos, o en términos más amplios por cualquier hecho, fenómeno o acontecimiento

Crterios para identificar a la población objeto de estudio

- **¿A qué o quiénes deseas ayudar a mejorar su condición?** Se refiere a la definición conceptual. Ejemplo de población amplia: postmenopáusicas, ejemplo específico: postmenopáusicas que reciben su terapia de reemplazo hormonal.
- **¿Dónde se encuentran estas unidades de estudio?** Se refiere a la delimitación espacial, amplio: la ciudad de Lima, específico: el cercado de la ciudad de Lima.
- **¿A qué periodo de tiempo corresponderá el estudio?** Se refiere a la delimitación temporal, amplio: 2026, específico: de enero a marzo del 2026. Ejemplo de población objeto de estudio: postmenopáusicas de la Ciudad de Lima en 2026.

Crterios para validar a la población objeto de estudio

- **¿Qué relación tienes con esta población?** Ejemplos de respuesta: son mis pacientes, son mis alumnos o son mis clientes.
- **¿Cómo accedes a esta población?** Si son mis pacientes: a la hora de la consulta, si son mis alumnos: a la hora de clases y si son mis clientes: a la hora de brindarles un servicio.
- **¿Cuál es tu responsabilidad con esta población?** Si son mis pacientes: a recuperar su salud, si son mis alumnos: a mejorar su rendimiento académico y si son mis clientes: a solucionar un problema a través de la prestación de servicios.

Criterios para delimitar a la población objeto de estudio

Además del lugar y el tiempo, se consideran los criterios de inclusión: para delimitar mejor a la población de estudio y los criterios de exclusión: para eliminar variables de confusión. Solo el resultado final de haber aplicado a los criterios de elegibilidad, se denomina **población objeto de estudio**. Incluso si se debe tomar una muestra, se hará sobre este resultado final.

Una población de estudio está conformada por unidades de estudio que habitualmente son confundidas con: las unidades de información, las unidades de observación, las unidades de análisis, las unidades de muestreo e incluso las unidades de experimentación.

La unidad de estudio es la entidad motivo de estudio y se refiere al qué o quién es sujeto de interés en una investigación. Es la unidad de la cual se necesita información o específicamente el dato. En las ciencias de la salud y las ciencias sociales, por lo general, la unidad de estudio es el individuo.

La unidad de información, los datos requeridos de la unidad de estudio no siempre, se obtienen de manera directa a partir de las unidades de estudio. Los datos primarios se obtienen de forma directa o indirecta, mientras que los datos secundarios solo de manera indirecta. La unidad de información es la entidad que nos brinda los datos de la unidad de estudio.

Unidad de observación, en un centro hospitalario los atendidos son las unidades de estudio, porque sobre ellos se requiere lograr un beneficio, pero las atenciones son las unidades de observación. De tal modo que un mismo paciente, puede tener varias atenciones, lo cual nos lleva a concluir que, una misma unidad de estudio, puede tener varias observaciones.

Unidad de análisis es la mínima cantidad de sustancia que se requiere para la medición de la variable analizada, esta cantidad es definida por el investigador para realizar sus mediciones. Por ejemplo, si queremos conocer el nivel de glucosa en plasma de un paciente diabético, 5 cc de sangre serán suficientes.

Unidad de muestreo, este concepto aparece únicamente cuando, para estudiar a una población, se requiere utilizar una muestral. Una muestra es un subconjunto de la población, una parte una fracción y los elementos que la conformen deben ser elegidos aleatoriamente. Por lo tanto, la unidad de muestreo es la unidad que se somete al proceso de aleatorización.

Unidad de experimentación, es el evento que recibe un estímulo y en el cual se observan los efectos de la intervención del investigador. Cada medición refleja la respuesta generada por dicha manipulación. Por ejemplo, si se busca reducir la duración del trabajo de parto, el evento estudiado es el propio parto, que ofrece una única oportunidad para modificar su duración.

El enunciado del estudio

Conocido coloquialmente en el ámbito académico como **el título de la tesis**.

Se trata de **una oración** conformada por:

- Un propósito del estudio (ejemplo: prevalencia) +
- Una línea de investigación (ejemplo: osteoporosis) +
- Una población de estudio (ejemplo: postmenopáusicas)

El resultado es: prevalencia de osteoporosis en postmenopáusicas.

Esto corresponde a: el enunciado del estudio.

Propósito, línea y población, son los elementos que conforman el enunciado del estudio, siendo el tercer elemento: población de estudio delimitado con lugar y tiempo para los estudios descriptivos

Quedando nuestro ejemplo, de la siguiente manera:

Prevalencia de osteoporosis en postmenopáusicas en Lima 2026.

Este ejemplo corresponde a un estudio descriptivo, que tiene una sensible diferencia, en cuanto al enunciado respecto del estudio analítico.

En ambos casos, estudio descriptivo y analítico, se conserva: el propósito del estudio, la línea de investigación y la población de estudio.

Siendo que, en el estudio descriptivo: la población es delimitada por lugar y tiempo, cuestión que no ocurre, en los estudios analíticos, en su lugar hace su aparición otra variable, y esto es porque los estudios analíticos relacionan variables.

La diferencia del enunciado de un estudio descriptivo y analítico se debe a que:

- Los estudios descriptivos buscan únicamente **validez interna**
- Los estudios analíticos además buscan **validez externa**.

La validez interna, se refiere a la validez de las conclusiones respecto de una población, incluso en los casos en que se extrae una muestra y se espera que los resultados y conclusiones obtenidos a partir de la muestra, sean también los que ocurran en la población, a este procedimiento se le denomina **inferencia** y es la garantía para alcanzar la validez interna, esto ocurre en los estudios descriptivos.

La validez externa, es una validez de extrapolación de las conclusiones, de una población estudiada, hacia otra que, no formó parte de la investigación. No tiene sentido entonces delimitar: el lugar, ni tampoco el tiempo, esto ocurre en los estudios analíticos.

Los estudios descriptivos buscan validez interna y por eso delimitan a su población de estudio, en cambio los estudios analíticos buscan validez externa, donde el lugar y tiempo únicamente corresponde a su ámbito de recolección de datos.

En los estudios descriptivos el algoritmo requiere de una delimitación espacial y temporal, para establecer el alcance de sus conclusiones, dado que los estudios descriptivos buscan únicamente validez interna. Por ejemplo: prevalencia de osteoporosis en postmenopáusicas en la Ciudad de Lima 2026. Los resultados serán válidos únicamente en Lima y para el 2026.

En los estudios analíticos el algoritmo requiere de la presencia de la otra variable analítica en el propio enunciado. Dado que los estudios analíticos, buscan además de la validez interna, la validez externa. Por ejemplo: asociación entre obesidad y diabetes en embarazadas, en este caso no es necesario, la delimitación espacial y temporal.

Para escribir tu mejor enunciado... Haz una lista de:

- 10 Líneas de investigación, dentro de tu campo del conocimiento.
- 10 Propósitos de estudio.
- 10 Poblaciones de estudio relacionadas con tu labor.

Por combinación, vas a obtener 1000 ideas de investigación o **1000 ideas de tesis**.

Todo trabajo de investigación es exhaustivo en cuanto a su intención

Si bien, el enunciado del estudio contiene al propósito del estudio, una sola oración puede ser insuficiente, para expresar la intención del investigador, para eso existe el planteamiento del estudio, un capítulo, una sección, en el que, se desarrolla la idea de investigación completa.

El enunciado del estudio no es sinónimo de problema de investigación

Así como el planteamiento del estudio, no equivale al planteamiento del problema. Porque incluso, cuando estás estudiando un problema, como línea de investigación, esto corresponde a: todo el recorrido, desde el descubrimiento del problema, hasta el planteamiento de su solución. En este momento estamos hablando únicamente de un estudio.

CAPÍTULO 3: EL MARCO TEÓRICO

Los antecedentes investigativos permiten identificar cómo ha sido abordado el fenómeno de estudio, qué resultados se han obtenido y cuáles son las limitaciones persistentes en la literatura científica. Su análisis crítico y ordenado sitúa el estudio dentro de la línea de investigación correspondiente, evitando duplicidades y orientando su enfoque metodológico. Asimismo, fundamentan la pertinencia y necesidad del nuevo trabajo al evidenciar el aporte específico que se pretende realizar al conocimiento acumulado.

El fundamento teórico agrupa las teorías, modelos y enfoques que explican el fenómeno estudiado, proporcionando coherencia conceptual al trabajo. Actúa como la base interpretativa que guía la comprensión y el análisis de los resultados. Su solidez es clave para garantizar la claridad, consistencia y profundidad del planteamiento investigativo.

Las definiciones conceptuales delimitan con precisión el significado teórico de los términos centrales del estudio, evitando ambigüedades y asegurando un lenguaje común. Se derivan directamente del fundamento teórico y expresan la esencia del constructo, contribuyendo a la coherencia y consistencia del marco teórico.

Las definiciones operacionales convierten los conceptos teóricos en indicadores observables y medibles, de acuerdo con el diseño del estudio. Precisan cómo se identificará empíricamente cada variable y garantizan la replicabilidad del procedimiento, constituyendo el vínculo fundamental entre la teoría y la medición.

El marco teórico

Es la estructura conceptual y argumentativa que sustenta, contextualiza y orienta una investigación, construida a partir de una **revisión rigurosa suficiente y pertinente de la literatura científica**.

Integra el estado del arte, los antecedentes, las teorías y las definiciones clave, organizados de forma coherente y jerarquizada, con el fin de identificar vacíos teóricos, fundamentar el estudio y garantizar una adecuada redacción académica.

Para una adecuada realización del marco teórico, **se requieren como mínimo los siguientes insumos**:

- El enunciado del estudio que contiene: la línea de investigación, el propósito del estudio y la población de estudio.
- El planteamiento del estudio que contiene: todos los elementos que justifican el estudio.
- El cuadro de variables que contiene: a la variable de estudio y las variables contextuales.

Los productos de esta labor clásicamente serán el:

- Marco conceptual o fundamento teórico.
- La definición operacional de términos.
- Los antecedentes investigativos.

Pasos para construir el marco teórico

Búsqueda de la literatura, selección de documentos, análisis de la información, organización del contenido y formato de redacción.

- Búsqueda de la literatura explora el conocimiento existente sobre el tema, identifica el estado del arte, enfoques, resultados y vacíos en investigaciones previas.
- Selección de los documentos: Filtra y elige fuentes suficientes y pertinentes actualizadas y confiables, alineadas con el propósito del estudio.
- Análisis de la información de la lectura de los documentos seleccionados: interpreta, compara y valora la información, identifica vacíos teóricos.
- Organización del contenido dispone el contenido de forma estructurada y jerarquizada para facilitar su comprensión.
- Formato de redacción aplica una estructura de presentación académica, clara, adecuada considerando normas de citación y estilo académico.

1. Búsqueda de la literatura

Consiste en realizar una búsqueda planificada sistemática y exhaustiva de fuentes académicas como: libros, artículos científicos, tesis, etcétera. Su objetivo, es conocer el estado del arte, identificar teorías y reconocer cómo se ha abordado el problema en investigaciones previas.

2. Selección de los documentos

Implica filtrar críticamente las fuentes académicas encontradas, para conservar solo aquellas que sean: suficientes que abarquen diversas perspectivas, pertinentes relacionadas con el problema en estudio y confiables de origen académico validado. Esta etapa, asegura una base sólida y relevante para la construcción del marco teórico.

3. Análisis de la información

Es un proceso de lectura crítica, interpretativa y comparativa de las fuentes seleccionadas, se busca identificar vacíos teóricos. El análisis se puede sistematizar mediante fichas matrices o esquemas, que faciliten la comprensión global del campo de estudio.

4. Organización del contenido

Consiste en estructurar la información de forma lógica, coherente y jerarquizada. Se puede organizar por temáticas, teorías, enfoques cronología, etcétera. Esta etapa determina el orden del marco teórico y prepara la base para una redacción clara y cohesionada.

5. Formato de redacción

Es la presentación final del marco teórico en formato académico, en general como: marco conceptual o fundamento teórico, definición operacional de términos y antecedentes investigativos. La redacción debe ser clara y formal respetando normas de citación y estilo

El marco teórico en un documento académico es un capítulo, no es así en el orden natural de las ideas, para el desarrollo de un trabajo de investigación.

La teoría está presente en todo momento, desde el planteamiento del estudio que responde a la pregunta ¿qué es lo que queremos investigar? hasta el planteamiento operacional que responde a la pregunta ¿cómo lo vamos a ejecutar? por esta razón, el planteamiento teórico es el marco que cubre a todo el proceso investigativo.

Marco conceptual o fundamento teórico

El marco conceptual es una sección del marco teórico, que se centra en definir delimitar y relacionar conceptualmente las categorías fundamentales del estudio.

Clásicamente dividimos al marco teórico en dos secciones: El marco conceptual o fundamento teórico y Los antecedentes investigativos.

En cualquier caso, el marco conceptual debe contener: además del fundamento teórico, a las definiciones conceptuales, definiciones operacionales y cuando corresponda a la postura epistemológica, entre otros, como la relación entre conceptos.

Componentes del marco conceptual:

Postura epistemológica es el enfoque filosófico relacionado al fenómeno en estudio. Fundamento teórico es la selección de teorías modelos y autores clave. Definiciones conceptuales contienen el significado teórico de los conceptos clave. Definiciones operacionales corresponden a la traducción empírica de los conceptos. Relación entre conceptos se refiere a las conexiones lógicas y explicativas entre conceptos.

1. Postura epistemológica

La postura epistemológica es la base filosófica, que orienta cómo se concibe la realidad (ontología), qué se reconoce como conocimiento válido (epistemología) y cómo se accede a él (metodología). Guía el enfoque teórico y metodológico de toda investigación. Ejemplos: positivismo, neopositivismo, constructivismo, fenomenología, hermenéutica, realismo crítico, etc.

2. Fundamento teórico

Corresponde al marco conceptual propiamente dicho y se refiere a la base conceptual que explica e interpreta el fenómeno en estudio, mediante teorías, modelos y enfoques relevantes articulando autores y postulados clave. Ejemplo: en la educación la teoría socio-constructivista de Vygotsky. En la salud el modelo de determinantes sociales de la salud de Dahlgren y Whitehead.

3. Definiciones conceptuales

Establecen el significado teórico de los conceptos clave de la investigación. Estas definiciones se derivan directamente del fundamento teórico y permiten aclarar la terminología para evitar ambigüedades. Ejemplo: la obesidad es el exceso de grasa corporal, que puede afectar la salud y aumentar el riesgo de enfermedades crónicas como diabetes y cardiovasculares.

4. Definiciones operacionales

Traducen las definiciones conceptuales, del lenguaje teórico a indicadores cuantificables u observables, útiles para su análisis y comparación.

Especifican ¿Cómo se medirán o identificarán empíricamente cada concepto teórico?

Ejemplo: la obesidad es, cuando el índice de masa corporal, calculado a partir del peso medido en kilogramos y estatura en metros al cuadrado, mayor o igual a 30 kg por metro al cuadrado.

5. Relación entre conceptos

Es la delimitación teórica de las conexiones asociativas, causales o funcionales entre las categorías centrales del estudio.

Sirve como base para plantear objetivos, hipótesis o construcción de modelos. Por ejemplo, la inactividad física se asocia con un mayor riesgo de obesidad, debido al menor gasto calórico diario.

Veamos un ejemplo del uso de los componentes del marco conceptual, para el estudio de la obesidad.

- **Postura epistemológica:** Se adopta un enfoque postpositivista que permite integrar métodos cuantitativos, para comprender la realidad social y sanitaria.
- **Fundamento teórico:** la obesidad se explica desde el modelo de determinantes sociales de la salud y teorías del comportamiento.
- **Definiciones conceptuales:** Se establecen los significados teóricos de términos clave como: obesidad, actividad física o nivel socioeconómico.
- **Definiciones operacionales:** Se traducen los conceptos a variables medibles mediante indicadores, como el índice de masa corporal para identificar la obesidad.
- **Relación entre conceptos:** La obesidad se vincula con factores conductuales sociales y ambientales en un modelo explicativo integral.

Los elementos que conforman el marco conceptual no necesariamente corresponden a su estructura. Las secciones o acápites del marco conceptual, inicialmente está guiadas por el fundamento teórico.

Ejemplo, en el modelo de determinantes sociales de la salud los acápites pueden ser: factores individuales, estilos de vida, ambiente social y condición socioeconómica.

Antecedentes Investigativos

Los antecedentes investigativos, son un conjunto estructurado de estudios previos, de naturaleza teórica o empírica, directamente relacionados con el problema de investigación.

Se componen esencialmente de textos académicos como: artículos científicos, tesis universitarias e informes técnicos, que presentan resultados ya obtenidos y metodologías aplicadas en el campo de estudio.

Funciones de los antecedentes investigativos

Son funciones que, comparten con el **marco conceptual**: delimitar el área de investigación, conocer el estado del arte, detectar vacíos del conocimiento, justificar la relevancia del estudio, orientar la metodología a desarrollar y servir como insumo para la discusión.

Siendo funciones exclusivas de los antecedentes investigativos:

- Evitar duplicación de estudios.
- Proporcionar bases empíricas para el desarrollo del estudio en curso.

1. Búsqueda de la literatura

La búsqueda de los antecedentes investigativos se realiza en función de la naturaleza como:

Antecedentes teóricos que se derivan de fuentes bibliográficas especializadas como libros académicos, artículos de revisión o ensayos especializados.

Antecedentes empíricos que provienen de investigaciones que reportan resultados concretos obtenidos mediante métodos sistemáticos, como artículos científicos, tesis universitarias y reportes técnicos.

2. Selección de documentos (criterios de selección)

- Calidad de la fuente, correspondiente al nivel de confiabilidad académica del documento consultado.
- Rigor metodológico relacionado a la solidez y coherencia en el diseño y ejecución del estudio.
- Relevancia temática, como el grado de vinculación directa con el problema de investigación.
- Vigencia científica, se refiere a la actualidad de la publicación en relación con el avance del conocimiento.
- Jerarquía lógica, correspondiente al orden estratégico de los antecedentes según alcance y contexto como: local, nacional e internacional.

3. Análisis de la información

La matriz de antecedentes investigativos, es una evolución académica de las antiguas fichas bibliográficas y corresponde a una herramienta metodológica que permite organizar, comparar y jerarquizar estudios previos.

Puede adoptar la forma de una tabla, con filas correspondientes a los estudios y columnas correspondientes a: autor y año, título del estudio, objetivo, metodología, principales hallazgos y pertinencia para el estudio.

La matriz de antecedentes investigativos, no se trata de un capítulo de la tesis, pero puede anexarse, si se desea documentar el proceso de **análisis crítico** de la información, evidenciando trazabilidad del trabajo realizado.

4. Organización del contenido (orden)

Se refiere al orden de presentación y puede ser:

- Por alcance geográfico: internacional, nacional y local.
- Por cronología: ascendente (de lo más antiguo a lo más reciente), descendente (de lo más reciente a lo más antiguo).
- Por nivel investigativo: de estudios descriptivos a explicativos según pertinencia.
- Por enfoque temático: agrupando las variables o categorías conceptuales del estudio.

5. Formato de redacción

La redacción debe emplear tercera persona y tiempo pasado utilizando un vocabulario científico, caracterizado por su formalidad, claridad y objetividad.

Cada párrafo debe presentar, autor o autores, el título del estudio, el año y lugar de realización si corresponde y su respectiva cita. Se sugiere seguir la **estructura OMRC** correspondiente a objetivo, metodología, resultados y conclusiones.

Finalmente, citar conforme al estilo de referenciación requerido como: APA, Vancouver u otro según norma institucional.

La construcción de la sección de antecedentes investigativos sigue los mismos pasos que para el marco conceptual: búsqueda, selección, análisis, organización y redacción.

Después de la búsqueda y selección, conviene realizar una **matriz de antecedentes** donde podemos observar un mapa completo de la información, que disponemos para su posterior organización y redacción.

CAPÍTULO 4: LAS VARIABLES

Las dimensiones de una variable permiten identificar las magnitudes físicas o dimensiones lógicas que la componen, según el modo en que serán medidas. Una variable puede ser unidimensional cuando requiere un solo indicador, o multidimensional cuando demanda varios indicadores. Esta estructura define la complejidad del concepto y orienta su estrategia de medición.

Las escalas de medición determinan la naturaleza del dato obtenido y los atributos que conserva cada nivel. Nominal y ordinal conforman las escalas categóricas; intervalo y razón constituyen las numéricas. Cada escala preserva atributos distintos orden, distancia u origen y condiciona el análisis estadístico que podrá aplicarse.

La relación entre variables expresa la función que cada una cumple dentro del nivel investigativo. En el nivel descriptivo, las variables se clasifican como variables de caracterización y variable de interés y; en el nivel relacional, como variables asociativas y de supervisión; en el nivel explicativo, como variables independientes y dependiente; en el nivel predictivo, como variables predictoras y variable a predecir; y en el nivel aplicativo, como variables de calibración y variable evaluativa.

La operacionalización transforma las variables en indicadores precisos que permiten obtener mediciones replicables. Implica seleccionar o proponer el indicador, definir unidades de medida y anticipar el tipo de variable resultante. Constituye el puente entre teoría y técnica, asegurando que cada variable pueda ser analizada con objetividad y coherencia metodológica.

Dimensiones de las variables

El concepto de dimensión aparece **cuando se requiere la medición**, para que las variables sean consideradas como tal, deben ser medibles.

En investigación, **las variables se definen como:** propiedades características o atributos observables en las unidades de estudio, recibe el nombre de variable, porque **expresa variabilidad**, es decir la dispersión de los valores finales de medición y no porque cambie ejemplo: el país de nacimiento.

La edad, el peso y la talla son ejemplos de variables correspondientes a **magnitudes físicas** como el tiempo, la masa o la longitud, que no son más que sustantivos abstractos, que se convierten en variables, cuando hacen referencia a un individuo, a una unidad de estudio.

Pero también, es posible evaluar en un individuo: la calidad de vida, la satisfacción del usuario y el rendimiento académico, en este caso correspondientes a **dimensiones lógicas**, son propiedades subyacentes que se supone poseen las unidades de estudio, individuos

A efectos de la medición **el concepto de dimensión es equivalente al de magnitud** que se evidencian a través de los indicadores...

Definición de indicador: es la forma de medir la variable, es la estrategia del investigador, para conocer el valor final de medición. Siendo así la naturaleza de las variables va a depender de la naturaleza del indicador seleccionado.

Tenemos:

- **Indicadores subjetivos** correspondientes a dimensiones lógicas que van a generar variables subjetivas y también tenemos...
- **Indicadores objetivos** correspondientes a magnitudes físicas que van a generar variables objetivas.

Ya sea que se traten de variables subjetivas u objetivas, estas pueden ser indistintamente unidimensionales o multidimensionales, dependiendo de la forma de realizar la medición.

Ejemplo de **variable objetiva unidimensional**

El peso, para conocer el valor final de medición del peso, solo hace falta el propio peso, algunos dicen que es un indicador directo, otros dicen que es su propio indicador. En cualquier caso, se trata de una magnitud física fundamental.

Ejemplo de **variable objetiva multidimensional**

El índice de masa corporal, para conocer el valor final de medición del índice de masa corporal se requiere de: el peso y la talla, claramente es una magnitud física derivada, el algoritmo: peso sobre talla al cuadrado, es el indicador y sus componentes son las dimensiones o magnitudes

Ejemplo de **variable subjetiva unidimensional**

El dolor, para conocer el valor final de medición del dolor de un paciente, aplicamos: la escala visual análoga, de una sola dimensión, la cual consiste en preguntar al paciente: en la escala del 0 al 10 ¿Cómo cuánto te duele? siendo cero la ausencia del dolor y 10 el dolor intolerable.

Ejemplo de **variable subjetiva multidimensional**

La calidad de la atención, para conocer el valor final de la calidad de la atención, habrá que aplicar un instrumento, el cual está compuesto por dimensiones.

Según Parasuraman, la calidad de la atención se mide a partir de las siguientes dimensiones: elementos tangibles, fiabilidad, capacidad de respuesta, seguridad y empatía siendo estos elementos las dimensiones que se requieren medir a través del indicador llamado cuestionario SERVQUAL.

Las variables se miden a través de sus indicadores que pueden ser

- Subjetivas correspondientes a **dimensiones lógicas** o indicadores
- Objetivas correspondientes a **magnitudes físicas**.

Las dimensiones lógicas se evidencian aplicando **instrumentos documentales**, mientras que las magnitudes físicas se evidencian aplicando **instrumentos mecánicos**.

Estas son las dimensiones de las variables: subjetivas, objetivas, unidimensionales y multidimensionales.

Escalas de medición de las variables

Las variables son propiedades características o atributos observables en las unidades de estudio, el resultado de su medición son los datos.

En estadística se les conoce como: **valores finales de medición** en los que habremos de analizar sus atributos, para diferenciar...

Cuatro escalas de medición: nominal, ordinal, de intervalo y de razón.

- Las escalas nominal y ordinal se agrupan como **variables categóricas** y reciben este nombre porque sus valores finales son categorías: masculino, casado, secundaria, severo, son algunos ejemplos.
- Las escalas de intervalo y de razón se agrupan como **variables numéricas** y reciben este nombre porque sus valores finales son números: 10 grados centígrados, 20 dólares, 80 kg, 3 hijos.

Las escalas nominal y ordinal se diferencian por el **atributo orden**.

Primaria, secundaria, superior, son las categorías ordenadas de la variable: nivel de instrucción.

Situación que, no ocurre con la variable: estado civil donde, soltero, casado o conviviente, no tienen ni guardan un orden en particular.

Las escalas ordinal y de intervalo se diferencian por el **atributo distancia**.

Es posible medir la distancia entre \$10 y \$20 que, equivale a la distancia entre \$20 y \$30, situación que no se puede conseguir, con una variable en escala ordinal, quien presenta un dolor de intensidad moderada no equivale al doble del dolor de intensidad leve.

Las escalas de intervalo y de razón se diferencian por el **atributo origen**.

Dado que, en la escala de razón contamos con un origen real, donde el cero significa la ausencia del individuo de la unidad de estudio.

No hay una persona que pese 0 kg, no hay un padre de familia que tenga cero hijos, en cambio la temperatura puede estar por debajo de cero grados centígrados, así como el saldo bancario, cuando nos cobran una comisión de mantenimiento

Las variables categóricas ya sean nominales u ordinales pueden ser dicotómicas si tienen solamente dos categorías o politómicas si tienen más de dos categorías.

Algunas personas piensan que las variables dicotómicas, no pueden ser ordinales, aquí algunos ejemplos de variable dicotómica ordinal: nuevo y continuador, para ser continuador en un servicio de salud antes tendrías que ser nuevo y por analogía: vivo - fallecido, sano - enfermo.

Las variables numéricas que agrupan a las escalas de intervalo y razón pueden ser continuas si provienen de medir y discretas si provienen de contar.

Las variables continuas, se pueden expresar con números enteros o fraccionarios, entre dos valores reales, siempre habrá un número intermedio real, mientras que las variables discretas solamente se pueden expresar con números enteros.

- La escala **nominal** posee categorías a las que se le asigna un nombre sin que exista ningún orden expreso entre ellas.
- Las variables **ordinales** poseen categorías ordenadas, pero no permite cuantificar la distancia entre una categoría y otra.
- Las variables **de intervalo** tienen intervalos medibles entre valores, al no tener un origen real pueden asumir valores negativos.
- Las variables en escala de **razón** tienen intervalos medibles con un origen real donde el cero significa la ausencia del individuo

Es posible convertir a una variable en escala de razón a escala de intervalo haciéndole **perder el atributo origen** como cuando convertimos la temperatura de la escala Kelvin a la escala centígrada.

También es posible convertir una variable en escala de intervalo a ordinal haciéndole **perder el atributo distancia**, como cuando convertimos la edad en categorías de: niño, adolescente, adulto y...

También es posible convertir a una variable en escala ordinal a nominal haciéndole **perder el atributo orden**.

En términos generales una variable numérica se puede convertir en categórica pero nunca al revés. Esto porque, podemos retirar atributos a los valores finales de una escala de medición, lo que no podemos hacer, es generar atributos que no poseen.

Relación entre las variables

Una línea de investigación se inicia con el descubrimiento de un problema y termina con el planteamiento de su solución, para ello hay que recorrer niveles, transitar, investigando a la variable de estudio, aquella que se obtiene por derivación de la línea de investigación.

Existen **líneas de investigación** que corresponden a:

Condiciones problemáticas

Como la diabetes en cuyo caso utilizaremos como variable de estudio a la diabetes con sus categorías presencia y ausencia, pero también podemos emplear como variable de estudio a la glucosa en ayunas con sus unidades miligramos por decilitro.

Otro ejemplo sería la hipertensión arterial en el que utilizaremos como variable de estudio a la propia hipertensión arterial con sus categorías presencia y ausencia, pero también podemos utilizar como variable de estudio a la presión arterial sistólica medida en milímetros de mercurio.

Por otro lado, existen **líneas de investigación** que corresponden a:

Situaciones que no son problemas

Como el rendimiento académico en ese caso podemos utilizar como variable de estudio al propio rendimiento académico en sus unidades de medición correspondientes o problematizar a la línea de investigación y utilizar como variable de estudio al bajo rendimiento académico, con sus categorías presencia y ausencia.

Finalmente, la rentabilidad de una empresa tampoco es un problema y se puede estudiar a la rentabilidad propiamente dicha o utilizar como variable de estudio a la buena rentabilidad con sus categorías presencia y ausencia

Si bien aquí hemos visto dos ejemplos de variable de estudio, por cada línea de investigación, puede haber más opciones. Elegir una u otra, depende del investigador, en función a la intención expresada en el propósito del estudio.

De la línea de investigación se deriva la variable de estudio y esa es una de las funciones del nivel exploratorio definir a la línea de investigación y también a la variable de estudio.

La variable de estudio tomará distintos nombres en los diferentes niveles de investigación, no porque correspondan a otra línea de investigación, sino por la función que cumplen de la relación entre las variables.

1. El nivel exploratorio únicamente puede operacionalizar a la variable de estudio, si el alcance de sus objetivos es definir operacionalmente, por otro lado, identificar, interpretar y determinar son el origen de la línea de investigación y de la construcción de la **variable de estudio**.
2. En el nivel descriptivo la variable de estudio toma el nombre de **variable de interés** y las otras variables se llaman **variables de caracterización**.

Se trata del nivel del análisis estadístico univariado, donde vamos a analizar a las variables de una en una, tanto a las de caracterización como a la de interés

3. El nivel relacional recibe este nombre precisamente porque relaciona a las variables y las hace de dos en dos por eso le dicen bivariado.

Es decir que la **variable de supervisión** que corresponde a la variable de estudio se va a relacionar con cada una de las **variables asociativas**, pero siempre de dos en dos y nunca vamos a relacionar a dos variables asociativas.

4. El nivel explicativo pretende demostrar una relación causa efecto, con el análisis estadístico multivariado, como uno de los criterios de causalidad.

Es decir, todas las **variables independientes** al mismo tiempo de forma multivariada, respecto de la **variable dependiente**, esto es un modelo multivariado.

5. En el nivel predictivo ya sabemos que existe una relación causal, entre las **variables exógenas** o predictoras respecto de la **variable endógena** o variable a predecir,

La intención aquí es crear un modelo predictivo para conocer lo que ocurriría, en caso de tener diferentes valores para mis variables predictoras. No se trata de una prueba estadística, sino de una técnica estadística para la construcción de un modelo predictivo.

6. En el nivel aplicativo sabemos cómo las **variables de calibración** influyen **sobre la variable evaluativa** y eso vamos a hacer precisamente, para obtener un resultado esperado vamos a cambiar deliberadamente los valores de las variables de calibración.

Operacionalización de las variables

Operacionalizar en términos simples se refiere a: **identificar, elegir o proponer el indicador** para obtener mediciones replicables que garanticen la objetividad del análisis estadístico y...

De repetir el estudio en una y otra ocasión conseguir resultados consistentes estando orientada al análisis estadístico la operacionalización es propia exclusiva y única de la investigación cuantitativa.

Las variables pueden tener indicadores objetivos o subjetivos. Si los **indicadores son objetivos** nos entregan datos numéricos o categóricos, mientras que si los **indicadores son subjetivos** los datos siempre son categóricos o cualitativos.

- Ejemplo de variable objetiva numérica: la edad identificamos el indicador fecha de nacimiento.
- Ejemplo de variable objetiva categórica: el sexo elegimos el indicador en recién nacidos caracteres sexuales primarios.

Por otro lado, cuando los indicadores de la variable son subjetivos, puede que existan instrumentos para su medición, o puede que tales instrumentos no existan

- Si existen instrumentos elegimos el indicador, ejemplo de variable: calidad de la atención, elegimos el cuestionario SERVQUAL.
- Pero qué pasaría si tales instrumentos no existen, proponemos el indicador y nos adentramos al campo de la creación y validación de instrumentos.

- Indicador es sinónimo de instrumento cuando la variable es subjetiva -

Operacionalicemos:

Haz un listado de **variables** participantes en el estudio, para luego identificar, seleccionar o proponer **indicadores**, a partir de los cuales conoceremos sus **valores finales de medición** e identificaremos el **tipo de variable** avizorando el análisis estadístico.

- Ejemplo de variable: peso, el indicador su propio peso.
- Ejemplo de variable: índice de masa corporal, el indicador es el peso y la talla al cuadrado.
- Ejemplo de variable: dolor, el indicador es la escala visual análoga.
- Ejemplo de variable: calidad de la atención, el indicador es el cuestionario SERVQUAL.

Kilogramos, son las unidades de los valores finales de medición de la variable peso, se trata de una variable numérica continua, porque proviene de medir y sus valores finales, se pueden representar incluso con cifras decimales.

Uno, dos y tres, son las unidades de los valores finales de medición de la variable paridad en, esta ocasión la variable numérica es discreta, porque proviene de contar y sus valores finales, únicamente pueden ser representados con números enteros.

Hacemos una diferenciación entre las variables numéricas continuas y discretas, por el tipo de distribución que pueden presentar. Las variables continuas presentan distribución normal, distribución t de Student; mientras que, las variables discretas presentan distribución binomial, distribución Poisson.

Leve, moderado y severo son los valores finales de medición de la variable: dolor, medida por la escala visual análoga, es una variable categórica ordinal.

Satisfecho e insatisfecho son los valores finales de la variable calidad de la atención, medida por el cuestionario SERVQUAL, se trata de una variable categórica nominal.

En las variables categóricas hacemos la diferencia por las escalas de medición.

En los estudios descriptivos, si bien el análisis estadístico es univariado agrupamos a las variables en variables de caracterización y variable de interés (variable de estudio). En el nivel relacional cada una de las variables que se van a relacionar con la variable de supervisión se llaman variables asociativas.

En el nivel explicativo agrupamos a las variables independientes para verificar su efecto o resultado sobre la variable dependiente. En el nivel predictivo hacemos un listado de variables exógenas o predictoras para conocer o predecir el valor de una variable endógena. En el nivel aplicativo las variables de calibración pretenden modificarse deliberadamente para obtener un resultado deseado sobre una variable evaluativa

Para realizar el cuadro de operacionalización de las variables, se requiere conocer tres conceptos:

- **Dimensionalidad** porque detrás de los indicadores objetivos están las magnitudes físicas, detrás de los indicadores subjetivos están las dimensiones lógicas
- **Tipo de variable** hacemos una diferencia en las variables numéricas entre continuas y discretas y, en las variables categóricas entre ordinales y nominales.
- **Relación entre variables** según el nivel de investigación.

CAPÍTULO 5: LOS OBJETIVOS

Los objetivos estructuran la estrategia operativa del estudio y traducen la intención cognoscitiva en acciones verificables. Constituyen el punto de partida del planteamiento operacional, ya que delimitan lo que debe alcanzarse para responder al propósito investigativo. Su redacción exige claridad, precisión y correspondencia directa con el enunciado del estudio.

Los objetivos cualitativos se orientan al descubrimiento, comprensión e interpretación de fenómenos, sin pretensión de inferencia estadística. Acciones como identificar, interpretar, determinar o definir reflejan esta lógica exploratoria y fenomenológica. Su alcance es heurístico, centrado en el significado y la construcción conceptual.

Los objetivos cuantitativos buscan medir, comparar, asociar, demostrar o predecir mediante procedimientos estadísticos. Responden a niveles investigativos descriptivos, relacionales, explicativos, predictivos o aplicativos, según la intención analítica. Su formulación depende de la operacionalización de variables y del diseño metodológico correspondiente.

El objetivo principal deriva directamente del enunciado del estudio al agregarle un verbo en infinitivo que exprese la acción operativa. **Los objetivos secundarios** complementan y desagregan el objetivo principal, permitiendo ejecutar los pasos necesarios para alcanzarlo. Ambos conforman un sistema coherente que guía la metodología y ordena el análisis.

Los objetivos del estudio

Cuando hablamos de objetivos, lo hacemos siempre en plural, porque estamos haciendo referencia al objetivo principal y a sus objetivos secundarios, que en conjunto se denominan: sistema de objetivos.

Pero antes... Hay que conocer dos conceptos: el primero origen de los objetivos, el segundo objetivos según el nivel investigativo.

Origen de los objetivos

Para poder plantear el objetivo principal basta con agregar al enunciado del estudio un verbo en infinitivo.

Veamos un ejemplo: en el enunciado del estudio “prevalencia de osteoporosis en posmenopáusicas” le agregamos un verbo en infinitivo “**estimar**”.

Quedando de la siguiente manera:

“**Estimar** la prevalencia de osteoporosis en postmenopáusicas”

Correspondiente al objetivo principal.

En principio, todo objetivo debe ser específico medible y alcanzable:

- **Específico** a partir del propósito del estudio prevalencia heredado de: el enunciado del estudio.
- **Medible** a partir de la operacionalización de la variable osteoporosis.
- **Alcanzable** a partir del verbo, estimar que, corresponde a una acción que se puede completar.

Veamos un segundo ejemplo, en el enunciado “factores de riesgo para el cáncer de pulmón” le agregamos un verbo en infinitivo “**determinar** los factores de riesgo para el cáncer de pulmón” es como quedaría el objetivo principal.

Finalmente, en el enunciado “tiempo de vida media en pacientes con cáncer de estómago” le agregamos un verbo en infinitivo “**pronosticar** el tiempo de vida media en pacientes con cáncer de estómago” así quedaría el objetivo principal.

Pero el objetivo **también puede ser una traducción operativa de: el Propósito del estudio:**

Comparación es un sustantivo abstracto, mientras que **comparar** es una acción que se puede completar.

Comparación y comparar son términos relacionados: Comparar es una traducción operativa de la comparación, mientras que comparación es una derivación nominal de comparar.

Veamos el siguiente ejemplo, en el enunciado: “comparación del rendimiento académico de dos salones de clases” traducimos... “**comparar** el rendimiento académico de dos salones de clases” así quedaría el objetivo principal.

Veamos un segundo ejemplo, en el enunciado “asociación de la obesidad con la diabetes en personas mayores de 40 años” traducimos... “**asociar** la obesidad con la diabetes en personas mayores de 40 años” de este modo quedaría el objetivo principal.

Finalmente, en el enunciado “correlación de la hemoglobina de la madre con el peso de su recién nacido” traducimos “**correlacionar** la hemoglobina de la madre con el peso de su recién nacido” quedándonos finalmente con el objetivo principal.

El enunciado del estudio no es igual no es equivalente, no es sinónimo del objetivo principal.

El objetivo principal es una derivación del **enunciado del estudio**, el enunciado del estudio responde a la pregunta ¿Qué es lo que queremos investigar? correspondiente a la idea del investigador. Mientras que el objetivo principal responde a la pregunta ¿Cómo lo vamos a ejecutar? correspondiente a la estrategia del investigador que, si bien puede cambiar, la idea de investigación nunca cambia.

Si desarrollamos el **enunciado del estudio**, lo que obtenemos es el **planteamiento del estudio** o planteamiento del problema. Mientras que si desarrollamos toda la estrategia a partir del **objetivo principal** lo que encontramos es el **planteamiento operacional**.

El enunciado del estudio es la expresión de la idea de investigación mientras que el objetivo principal es la estrategia para alcanzar el propósito que encontramos en el enunciado del estudio, ya sea que tengamos que agregar un verbo en infinitivo al enunciado del estudio o traducir al propósito mismo, en ambas situaciones nos estamos refiriendo a una operacionalización del enunciado del estudio en objetivo principal.

Todos los manuales de investigación científica coinciden en que **los objetivos comienzan con un verbo en infinitivo**, pero no se trata de cualquier verbo, sino de aquellos que podemos alcanzar según el nivel investigativo, lo cual veremos en nuestra siguiente presentación.

Objetivo según el nivel investigativo

Nivel exploratorio

- **Identificar** se encarga de evidenciar el fenómeno hecho o acontecimiento, por ejemplo: evidenciar la presencia de la enfermedad, aun cuando no se le entienda, su principal herramienta: la fenomenología.
- **Interpretar** permite argumentar un fenómeno o por lo menos tratar de entenderlo, para ello haremos uso de la hermenéutica, esto se puede entender como el arte de explicar, traducir o esclarecer los hechos.
- **Precisar** intenta conseguir una secuencia lógica de procedimientos para detectar la presencia del fenómeno, en el campo de la salud, esto equivale a diagnosticar, haciendo uso del razonamiento heurístico.
- **Definir** se ocupa de conceptualizar y operacionalizar el fenómeno con el fin de desarrollar una línea de investigación, de modo que dos o más investigadores puedan coincidir, su principal herramienta el constructivismo.

Nivel descriptivo

- **Caracterizar** se ocupa de describir las características de un solo grupo que posee una determinada condición, esta condición puede ser un problema una enfermedad etcétera, aparece aquí una nueva herramienta: la estadística.
- **Estimar** se encarga de calcular un parámetro en la población, a partir de una muestra, esta necesidad se presenta cuando no es posible estudiar a toda la población, aparece una nueva herramienta: la inferencia.
- **Verificar** se trata de comparar o contrastar los descriptivos de un grupo de estudio, respecto de los parámetros de una población, aquí hace su aparición el ritual de la significancia estadística.

Nivel relacional

- **Comparar**, hay dos formas de comparar, se puede comparar dos grupos paralelos de personas, pero también se puede comparar dos medidas repetidas en el mismo grupo de personas. aparece aquí el análisis bivariado.
- **Asociar** es análogo al objetivo correlacionar, el primero se refiere a la concurrencia de dos sucesos, el segundo a la correlación de unidades, participan en ambos casos dos variables aleatorias.
- **Medir la fuerza de asociación** la fuerza de asociación o la fuerza de correlación se ejecuta, cuando se ha demostrado previamente la existencia de asociación o correlación respectivamente, se utilizan aquí los índices y coeficientes.

Nivel explicativo

- **Evidenciar** busca revelar la relación causal, en un estudio observacional, no es un experimento, no hay variable manipulada requiere de criterios de causalidad, más allá de la estadística, por ejemplo, la relación temporal.
- **Demostrar** requiere la intervención del investigador para manipular la variable independiente y demostrar sus efectos sobre la variable dependiente, se trata de un experimento que, cuenta con manipulación y control.
- **Probar** corresponde a la posibilidad de conseguir resultados consistentes en una y otra ocasión ya sea por el propio investigador o por parte de otros investigadores, se puede medir también la intensidad del efecto.

Nivel predictivo

- **Predecir** calcula la probabilidad de ocurrencia de un suceso en una serie de eventos, utiliza procedimientos como la regresión logística binaria o la regresión lineal múltiple, también utiliza la minería de datos.
- **Pronosticar** el análisis de supervivencia, permite calcular el tiempo medio en que ocurre un suceso el pronóstico se refiere a la estimación puntual con sus respectivos intervalos de confianza del tiempo de vida media.
- **Prever** aquí concurren varios objetivos prever, prevenir, proyectar, prever, el análisis de las series de tiempo permite conocer el comportamiento de una variable en función a sus propios datos del pasado.

Nivel aplicativo

- **Optimizar** se requiere de un monitoreo para evidenciar que la intervención ha conseguido pasar de un estado a otro mejor, lo que está en evaluación en este caso es el éxito de la intervención en términos cuantitativos.
- **Calibrar** mide el resultado de la intervención respecto de los recursos empleados, utiliza el análisis de la capacidad, para asegurar resultados dentro de los límites de normalidad o límites de especificación inferior y superior.
- **Aceptar** utiliza el muestreo de aceptación, para conocer el número de unidades necesarias, para validar un procedimiento proceso o tratamiento, determinando así la aceptación o validación del procedimiento.

Objetivo principal y objetivos secundarios

El objetivo principal busca completar la intención del investigador, expresada en el enunciado del estudio, mientras que los objetivos secundarios, buscan completar al objetivo principal. Todos los manuales de investigación señalan que las características de los objetivos son:

- **Alcanzable** representa la acción que se puede completar.
- **Específico** derivado del propósito del estudio o especificidad del estudio.
- **Medible** en virtud de la operacionalización de la variable de estudio.
- **Relevante** para el propósito del estudio en caso del objetivo principal.

El objetivo principal es relevante para el enunciado del estudio, mientras que los objetivos secundarios son relevantes para el objetivo principal.

Podemos presentar sinópticamente a los objetivos, en función a sus características:

1. **Específico:** El objetivo principal se deriva del propósito o especificidad del estudio, los objetivos secundarios son pasos intermedios, para alcanzar el objetivo principal.
2. **Medible:** El objetivo principal se deriva de la operacionalización de las variables, los objetivos secundarios exactamente igual.
3. **Alcanzable:** El objetivo principal nace de la operacionalización del propósito del estudio, es por esta razón, que al objetivo principal se le conoce también como la versión operativa del enunciado del estudio.

Los objetivos secundarios dan continuidad, a esta estrategia metodológica y estadística, iniciada en el paso anterior. El objetivo principal da inicio al planteamiento operacional de todo trabajo de investigación, es por esta razón que la propiedad de alcanzable es la más importante de todas

4. **Relevante:** El objetivo principal es relevante para el propósito del estudio, los objetivos secundarios son relevantes para el objetivo principal.

Sistemas de objetivos algorítmicos y heurísticos

Ejemplo de sistema algorítmico

Objetivo principal

- Comparar el promedio del rendimiento académico de dos salones de clases.

Objetivos secundarios

- Estimar el promedio del rendimiento académico del salón de clases A.
- Estimar el promedio del rendimiento académico del salón de clases B.

Y pudiera haber un tercer objetivo secundario, si es que así lo expresa el objetivo principal

Ejemplo de **sistema heurístico**

Objetivo principal

- Describir las características de los pacientes con la enfermedad de covid-19

Objetivos secundarios

- Describir los síntomas de los pacientes con la enfermedad de covid-19
- Describir los signos de los pacientes con la enfermedad de covid-19

Y si el investigador lo considera pertinente, puede agregar más objetivos secundarios como

- Describir los signos radiológicos de los pacientes con la enfermedad de covid-19

Diferencia entre un sistema algorítmico y un sistema heurístico

Correspondiente a los estudios inferenciales y no inferenciales respectivamente:

El sistema algorítmico es un conjunto de reglas definidas ordenadas y finitas para completar una intención es como una suma para alcanzar el objetivo principal, paso uno, paso dos, paso, tres.

El sistema heurístico por su lado es el arte y la ciencia del descubrimiento, para lo cual usa criterios principios y razonamientos, como en un diagnóstico médico: signos, síntomas y todos los exámenes auxiliares que considere el médico tratante.

Recordemos **los estudios no inferenciales** buscan concluir únicamente sobre el grupo estudiado, mientras que **los estudios inferenciales** buscan llevar las conclusiones del grupo estudiado a un conjunto mayor llamado población o incluso hacia otras poblaciones que, no son parte de su ámbito de recolección de datos.

CAPÍTULO 6: LA HIPÓTESIS

La formulación de hipótesis consiste en plantear una proposición verificable que anticipa un resultado esperado del estudio, ya sea a partir de la observación empírica directa o del sustento teórico disponible. Su redacción exige coherencia lógica y correspondencia con el nivel investigativo, orientando el diseño del estudio y, cuando corresponde, el análisis estadístico.

La prueba de hipótesis es el procedimiento mediante el cual se contrasta empíricamente la afirmación planteada, evaluando si la evidencia respalda o no la proposición. Este proceso requiere definir supuestos, parámetros y criterios de decisión. Su finalidad es determinar la consistencia entre los datos observados y la hipótesis formulada.

La prueba estadística es el procedimiento específico utilizado para contrastar una hipótesis y su elección debe realizarse considerando seis criterios fundamentales: el nivel investigativo, el objetivo estadístico, el tipo de variables involucradas, el tipo de estudio, el diseño del estudio y el comportamiento de los datos. Cada prueba se sustenta en supuestos técnicos particulares cuya adecuada verificación garantiza la validez de la inferencia, por lo que una selección correcta resulta determinante para la solidez de las conclusiones.

Los estudios sin hipótesis son aquellos en los que no se busca contrastar proposiciones previas, sino describir, caracterizar o explorar fenómenos de interés. Pueden incluir estimaciones puntuales de parámetros o medidas de efecto, como el odds ratio (OR) o el riesgo relativo (RR), sin recurrir a pruebas de contraste inferencial. Su finalidad es aproximarse al objeto de estudio de manera abierta, favoreciendo una comprensión del fenómeno.

La prueba de hipótesis

Una hipótesis es una proposición, una oración susceptible de asignarle uno de los **valores de verdad** de: verdadero que corresponde a la hipótesis alterna, o falso que corresponde a la hipótesis nula.

La hipótesis del investigador es la hipótesis alterna, es lo que el investigador desea demostrar y se expresa como una afirmación anticipada.

El ritual de la significancia estadística al servicio de la prueba de hipótesis, para los estudios cuantitativos, descrito originalmente por Fisher trae cinco pasos: planteamiento de hipótesis, nivel de significancia, prueba estadística, cálculo del p-valor y toma de decisiones.

1. Planteamiento de hipótesis

Se trata de escribir a la hipótesis nula y a la hipótesis del investigador. Desde **el punto de vista lógico**, las hipótesis tienen dos valores de verdad: verdadero y falso.

Siendo verdadero, la hipótesis alterna, conocida también como hipótesis del investigador y se le denota como H_1 . En los estudios comparativos la hipótesis alterna, es la que busca demostrar las diferencias, ya luego la hipótesis nula únicamente se opone a la hipótesis del investigador.

2. Nivel de significancia

Se le denota por la letra griega alfa (α) y se establece de forma convencional y preliminar en 5%, en cifras decimales 0,05 pudiendo ser mucho menos, dependiendo de la línea de investigación.

El nivel de significancia, es la máxima cantidad de **error tipo I**, que estamos dispuestos a: aceptar, para dar como válida la hipótesis del investigador (H_1).

Es en este momento que debemos recordar el concepto de error tipo I...

El error tipo I ocurre cuando aceptamos la hipótesis del investigador, cuando en realidad esta afirmación era falsa, es un juicio de valor equivocado.

A continuación, debemos cuantificar la magnitud de este error, y eso es precisamente **el p-valor**, corresponde a la probabilidad de equivocarse, al aceptar la hipótesis del investigador como verdadera, es la probabilidad de cometer un **error tipo I**.

Así, debemos identificar estos tres conceptos que, nos permitirán hacer la interpretación de nuestros resultados estadísticos.

3. Elegir una prueba estadística

Existen seis criterios para elegir adecuadamente un procedimiento estadístico y estos son:

- El nivel investigativo
- El objetivo estadístico
- El tipo de variables
- El tipo de estudio
- El diseño del estudio
- El comportamiento de los datos

(Criterios que desarrollaremos en la siguiente presentación).

4. Cálculo del p-valor

Se trata de cuantificar la magnitud del **error tipo I**, a partir del procedimiento estadístico elegido en el paso anterior.

Este paso, no fue planteado originalmente por **Fisher**, porque en esa época no se disponían de programas informáticos que, ayuden a calcular el p-valor, el cual, cuantifica el error tipo I, la probabilidad de equivocarse cuando el investigador hace una afirmación.

A día de hoy, esta probabilidad se calcula con cualquier software estadístico.

5. Regla de decisiones

Para saber, si nos quedamos con la hipótesis nula o la hipótesis del investigador, en función al contraste del p-valor respecto del nivel de significancia.

Todo investigador pretende quedarse con H_1 , porque se trata de su proposición.

Por tanto, espera que la magnitud del error tipo I, sea de la menor cantidad posible y de cuánto estamos hablando... de un número que esté por debajo del nivel de significancia (preliminarmente establecido).

Cuando el p-valor está por debajo del nivel de significancia, nos quedamos con H_1 (es la hipótesis del investigador). Pero si el p-valor está por encima del nivel de significancia preliminarmente establecido, nos quedamos con H_0 (es la hipótesis nula).

Esta regla de decisiones, sirve para utilizarla con cualquier prueba estadística, en cualquier trabajo de investigación.

La prueba estadística

Seis criterios para elegir una prueba estadística: nivel investigativo, objetivo estadístico, tipo de variables, tipo de estudio, diseño del estudio y comportamiento de los datos.

1. Nivel investigativo

El nivel exploratorio que si bien cuenta con datos, los analiza sin el uso de la estadística. El nivel descriptivo aparece el análisis estadístico univariado, el nivel relacional completa sus objetivos con el análisis estadístico bivariado, el nivel explicativo que busca demostrar relaciones de causalidad se basa en el análisis estadístico multivariado, el nivel predictivo usa modelamiento predictivo, finalmente el nivel aplicativo completa sus objetivos con las metodologías de mejora continua.

2. Objetivo estadístico

Como ejemplo nos vamos a ubicar en el nivel investigativo relacional:

- Objetivo correlacionar, prueba estadística: correlación de Pearson.
- Objetivo asociar, prueba estadística: chi cuadrado de Independencia.
- Objetivo comparar, prueba estadística: t de Student para muestras independientes.

Así, existe un procedimiento estadístico para cada uno de los objetivos en todos los niveles de la investigación cuantitativa.

3. Tipo de variable

Ejemplo: En el nivel relacional con dos variables analíticas:

- Si ambas variables son categóricas prueba estadística: chi cuadrado.
- Si ambas variables son numéricas prueba estadística: correlación de Pearson.
- Si una variable es categórica y la otra numérica: t de Student al caso de que, existan únicamente dos grupos.
- Si una variable es numérica y la otra categórica igual que en el ejemplo anterior.

4. Tipo de estudio

Continuamos con el nivel relacional, objetivo comparar, variable aleatoria numérica, estudio transversal de una sola medida, comparación entre grupos o comparación entre sujetos prueba estadística: t de Student para muestras independientes, al caso de que sean solamente dos grupos.

En el mismo nivel, con mismo objetivo, misma variable aleatoria numérica, estudio longitudinal, medidas repetidas, comparación intragrupo o comparación intra sujeto, la prueba estadística sería: t de Student para muestras relacionadas, al caso de que sean únicamente dos medidas.

5. Diseño del estudio

Se refiere a las características del estudio.

El diseño de cohortes es: prospectivo y longitudinal, mientras que el diseño de casos y controles es: retrospectivo y transversal, ambos se presentan en una tabla de 2 por 2.

- **En el diseño de cohortes** analizamos la incidencia de la enfermedad en la cohorte expuesta, para dividirlo entre la incidencia de la enfermedad en la cohorte no expuesta, el resultado se llama riesgo relativo (RR).
- **En el diseño de casos y controles**, ubicamos un grupo de casos, para buscar la frecuencia del factor de exposición y relacionarlo con la frecuencia del factor de exposición, respecto del grupo de los controles, el resultado se denomina Odds Ratio (OR).

Mismo objetivo misma tabla de 2 por 2, dos procedimientos estadísticos distintos, en relación al diseño del estudio.

6. Distribución de los datos

Las pruebas paramétricas se aplican a los datos numéricos continuos que, cuentan con distribución normal, si esto es así, cuando contrastamos la media de una muestra y dibujamos las barras de error que, contengan a la media poblacional, decimos que no hay diferencias, pero esto solamente ocurre cuando contamos con distribución normal.

Si los datos con los que estamos trabajando no cuentan con distribución normal, no se pueden aplicar procedimientos estadísticos paramétricos.

En los estudios transversales. Si contamos con un solo grupo y la variable aleatoria es numérica: t de Student para una muestra. Si son dos grupos: t de Student para muestras, independientes. Si son más de dos grupos: análisis de la varianza con un factor intersujetos.

En los estudios longitudinales. Si la variable aleatoria es numérica, dos medidas: t de Student para muestras relacionadas o medidas repetidas. Más de dos medidas: análisis de la varianza con un factor intrasujeto. Si los datos no cuentan con distribución normal: ocupamos su equivalente no paramétrico.

Tipos de hipótesis

1. Desde el punto de vista de la lógica proposicional

Las hipótesis tienen dos valores de verdad: verdadero y falso. Verdadero correspondiente a la hipótesis alterna (H_1) y falso correspondiente a la hipótesis nula (H_0).

En los estudios comparativos, la hipótesis alterna o hipótesis del investigador, busca descubrir las diferencias, ya luego la hipótesis nula únicamente se opone a la hipótesis del investigador.

- La hipótesis nula (H_0) es llamada también hipótesis de trabajo, corresponde al **valor de verdad de falso**, es lo que el investigador desea descartar.
- La hipótesis alterna (H_1) es llamada también hipótesis del investigador, corresponde al **valor de verdad verdadero**, es lo que el investigador desea demostrar.

Ejemplo:

- H_0 : El hábito de fumar **no** es un factor de riesgo para el cáncer de pulmón
- H_1 : El hábito de fumar **si** es un factor de riesgo para el cáncer de pulmón

2. Desde el punto de vista matemático

Las hipótesis tienen dos planteamientos de dos colas y de una sola cola.

En los estudios comparativos **de dos colas** la hipótesis del investigador plantea las diferencias, ya sea que el grupo uno, es menor que el grupo dos, o que el grupo uno sea mayor que el grupo dos

En los estudios comparativos **de una sola cola** la primera versión indica que, el grupo uno es menor al grupo dos y la segunda versión el grupo uno es mayor que el grupo dos, en todos los casos la hipótesis nula, se opone a la hipótesis del investigador.

En los estudios de **correlación de dos colas** la hipótesis del investigador plantea que el coeficiente de correlación “ r ” es distinto de cero ya sea que su valor sea negativo, correlación inversa o que su valor sea positivo correlación directa.

En los estudios de **correlación de una sola cola**: En su primera versión la hipótesis del investigador plantea que el coeficiente de correlación “ r ” es menor a cero, es negativo. En su segunda versión la hipótesis del investigador tiene un coeficiente de correlación “ r ” mayor a cero un valor positivo.

3. Desde el punto de vista gramatical

Las hipótesis tienen dos partes fundamento y deducción.

Ejemplo:

Dado que, el humo del cigarrillo contiene alquitrán benceno y otras sustancias cancerígenas...

Es probable, que el hábito de fumar sea un factor de riesgo para el cáncer de pulmón.

El fundamento suele comenzar con la frase “dado que” y se trata de un razonamiento argumentativo.

La deducción suele comenzar con la frase “es probable que” se trata de la hipótesis misma es H_1 .

El fundamento está basado en la evidencia científica de estudios previos y teorías existentes, es el razonamiento argumentativo que, nos lleva a sostener la afirmación propuesta, es parte de la justificación del estudio.

La deducción es el enunciado del estudio, calificado con el valor de verdad de verdadero, es la hipótesis alterna (H_1) es lo que el investigador desea demostrar, es la conclusión del proceso lógico que se puede verificar empíricamente.

4. Desde el punto de vista de su origen

Las hipótesis son empíricas y racionales

Las hipótesis empíricas se originan en la observación directa de la realidad, no tienen un fundamento teórico, se formulan a partir de las observaciones de los datos y de patrones empíricos, su origen es puramente fenomenológico.

Las hipótesis racionales exigen un fundamento, son propias de los estudios con intervención, se fundamentan en la evidencia científica, en conocimientos previos y en las teorías establecidas, su origen es teórico.

Una intervención implica la acción directa del investigador, sobre las unidades de estudio, en este caso, es esencial que esta intervención, esté respaldada por una hipótesis racional o un fundamento teórico, proporcionando así una base sólida, para justificar las acciones del investigador.

CAPÍTULO 7: EL MUESTREO

El cálculo del tamaño de muestra determina cuántas unidades deben observarse para obtener estimaciones precisas y válidas. Depende del nivel de confianza, nivel de precisión, variabilidad y tipo de variable o contraste requerido. Un tamaño insuficiente reduce la potencia; uno excesivo incrementa costos sin mejorar la validez.

El muestreo aleatorio se basa en un proceso probabilístico donde cada unidad tiene la misma oportunidad de ser seleccionada. Este método permite el uso legítimo de inferencia estadística y reduce la influencia de factores subjetivos. Su fortaleza radica en la representatividad y en la minimización de sesgos sistemáticos.

El muestreo no aleatorio selecciona unidades mediante criterios intencionados, de conveniencia o de accesibilidad, respondiendo a condiciones prácticas del estudio. Aunque limita la capacidad de generalización, puede ser pertinente en fases exploratorias o poblaciones difíciles de acceder. Su utilidad depende de la coherencia con el propósito investigativo.

Los sesgos de selección aparecen cuando la muestra difiere sistemáticamente de la población objetivo, afectando validez interna y externa. Pueden originarse por errores en el marco muestral, criterios inadecuados o no respuesta diferencial. Identificarlos y mitigarlos es esencial para garantizar conclusiones confiables.

Población y muestra

Una población es un conjunto de unidades de estudio, que son objeto de interés por parte del investigador.

Cuando no es posible acceder a toda la población estudiamos una parte, un segmento, una fracción la denominaremos **muestra**...

Y esperamos que, los resultados obtenidos en la muestra sean los mismos de haber estudiado a toda **la población**, para ello debemos extraerla mediante un procedimiento denominado **muestreo**.

Pero, no todos los estudios buscan estudiar poblaciones, incluso a partir de muestras, sino únicamente **los inferenciales** buscan estudiar a toda la población y cuando esto no es posible se opta por una muestra

Mientras que, los estudios **no inferenciales** buscan estudiar una o más unidades de estudio que no representan a una población.

Existen al menos tres situaciones en las que recurrimos al muestreo:

- Cuando la población es muy grande **inalcanzable**, por su magnitud.
- Cuando la población no cuenta con un marco muestral, **desconocida** en tamaño.
- Cuando la población es **inaccesible** para su completa examinación, o la examinación de las unidades de estudio representa su destrucción.

Para que una muestra sea representativa de la población, debe cumplir con dos condiciones:

- El cálculo del tamaño de la muestra a través de un algoritmo matemático.
- Las técnicas de muestreo o de selección.

El cálculo del tamaño de la muestra busca controlar el error aleatorio a través de un algoritmo matemático.

Las técnicas de muestreo o de selección buscan controlar el error sistemático a través del control de los sesgos de selección.

Error aleatorio y error sistemático que son amenazas de la validez de un estudio. Veamos un ejemplo del cálculo del tamaño de la muestra en cada uno de los niveles de investigación

1. Muestreo en el nivel exploratorio

Los estudios exploratorios no son inferenciales, no buscan extrapolar sus resultados hacia un conjunto mayor, por lo tanto no existe el cálculo del tamaño de la muestra y la técnica de selección será no aleatoria, por ejemplo muestreo por conveniencia o en bola de nieve.

2. Muestreo en el nivel descriptivo

Utiliza al **nivel de confianza** y a la precisión para el **cálculo del tamaño de la muestra**, ya sea para la estimación de frecuencias o promedios, con marco muestral conocido o desconocido aquí el tamaño de la muestra es el número mínimo de unidades de estudio que se deben incluir.

3. Muestreo en el nivel relacional

Aparece un segundo concepto, el de **potencia de prueba** relacionado al error tipo II de manera que el cálculo del tamaño de los grupos, es exacto no se puede aumentar ni se puede disminuir, además de que “n” minúscula representa al tamaño de cada grupo.

4. Muestreo en el nivel explicativo

Utiliza los mismos algoritmos que el nivel anterior, con la diferencia que, en los experimentos nos encontramos con muestras insuficientes, en cuyo caso si el p-valor es menor al nivel de significancia, el resultado sigue siendo concluyente e inferencial, mientras que si el p-valor es mayor o igual al nivel de significancia el resultado no es concluyente. “La ausencia de evidencia no es evidencia de ausencia”.

5. Muestreo en el nivel predictivo

El Muestreo, es el núcleo de los algoritmos genéticos, ya sea para la selección de padres, para la mutación o la evolución en la construcción de los algoritmos genéticos, en la minería de datos y la inteligencia artificial

6. Muestreo en el nivel aplicativo

La probabilidad de aceptación del lote, se calcula en función al tamaño de la muestra, el número máximo de defectos aceptables y la proporción de defectos en el lote. Para ello antes hay que definir los criterios de aceptación como el número máximo de defectos permitidos.

Si la variable analizada es categórica y se le denomina muestreo de aceptación por atributos o el rango de tolerancia definido si la variable analizada es numérica y se le denomina muestreo de aceptación por mediciones.

Muestreo aleatorio (probabilístico)

Para que una muestra sea representativa de su población el muestreo debe incluir al cálculo del tamaño de la muestra y a las técnicas de muestreo.

Las técnicas de muestreo pueden ser probabilísticas o no probabilísticas.

Los procedimientos probabilísticos, aseguran que los resultados del estudio sean más exactos, mientras que los procedimientos no probabilísticos, aseguran que la ejecución del estudio sea más factible.

Los muestreos aleatorios o probabilísticos son: simple, sistemático, estratificado y por conglomerados.

1. Muestreo aleatorio simple

Es un método de selección en el que, cada elemento de la población tiene la misma probabilidad de ser incluido en la muestra, para esto hay que definir el tamaño de la población “N” mayúscula y calcular el tamaño de la muestra “n” minúscula.

Para ello, debemos enumerar o enlistar a todos los elementos de la población del uno al “N” mayúscula y seleccionar “n” minúscula elementos de manera completamente aleatoria.

El muestreo aleatorio simple requiere de un **marco muestral** y que sus elementos sean accesibles, es el más representativo y simple de aplicar en poblaciones pequeñas, pero es poco práctico para poblaciones muy grandes y dispersas.

Cada vez que se quiere calcular un parámetro de la población a partir de una o más muestras, se obtienen resultados numéricos sensiblemente distintos, a estas diferencias, se les conoce como **error aleatorio** y será mucho menor Mientras más grande sea la muestra siempre que se acompañe del muestreo aleatorio simple.

2. Muestreo aleatorio sistemático

Es un método de selección que implica, elegir elementos de una población a **intervalos regulares** a partir de un punto de inicio seleccionado al azar. Al igual que en el caso anterior hay que definir el tamaño de la población “N” mayúscula y calcular el tamaño de la muestra “n” minúscula, calcular también el intervalo de selección o salto sistemático denotado por “k” donde “k” es igual a “N” mayúscula sobre “n” minúscula.

Luego elegir un punto de inicio, de manera aleatoria dentro del intervalo de 1 a “k” a partir de este punto de inicio seleccionar cada k-ésimo elemento.

El muestreo aleatorio sistemático asegura que cada elemento de la población tenga una probabilidad predefinida de ser incluido en la muestra. Para ello hay que verificar que no haya periodicidad en la población que coincida con el intervalo “k”.

Vamos a suponer que el salto sistemático “k” es igual a 11 elegimos un arranque aleatorio entre el 1 y el 11 si este resultado fuera el número 2, entonces comenzamos a construir la muestra entre el 2, 13, 24, 35, hasta completar “n” minúscula.

3. Muestreo aleatorio estratificado

Se usa cuando en la población se identifican estratos o **subgrupos homogéneos** y se seleccionan elementos al azar dentro cada estrato.

En este caso, hay que dividir a la población en estratos según su característica relevante, determinar el tamaño de la muestra “n” minúscula y su afijación de estratificación, que puede ser proporcional en relación al tamaño, de Neyman en relación a la variabilidad, u óptima en relación a los costos de medición finalmente realizar un muestreo aleatorio simple dentro de cada estrato.

Se usa cuando en la población existen subgrupos y se busca representarlos a todos. Los estratos deben ser completamente exhaustivos y mutuamente excluyentes, para conseguir maximizar la precisión enfocándose en los estratos más variables. La afijación más utilizada es la **afijación proporcional** aquella donde el tamaño de la muestra en cada estrato es proporcional al tamaño de la población en cada estrato.

4. Muestreo por conglomerados

Es un método de selección cuando en la población se identifican conglomerados o grupos heterogéneos que representan a toda la población. Aquí hay que dividir a la población en conglomerados exhaustivos y mutuamente excluyentes, seleccionar los conglomerados mediante muestreo aleatorio simple y elegir aleatoriamente los elementos dentro de los conglomerados seleccionados.

El muestreo aleatorio por conglomerados implica que, cada conglomerado debe reflejar las características de la población en su conjunto.

Es un muestreo de una etapa, si se estudian todos los elementos del conglomerado, pero es un muestreo en dos etapas si en cada conglomerado se elige un grupo. Llamado también muestreo por clúster, implica encontrar la heterogeneidad que se observa en la población en cada conglomerado o clúster, de manera que nos aboquemos en seleccionar únicamente estos conjuntos.

Muestreo no aleatorio (no probabilístico)

Clásicamente presentamos al muestreo no aleatorio en cuatro conjuntos: muestreo por cuotas, muestreo según criterio, muestreo en bola de nieve y muestreo por conveniencia.

Decimos grupos, porque dentro de cada uno de ellos **existen subtipos**, que algunos autores incluso los denominan como sinónimos.

Siendo el muestreo por cuotas aquel que exhibe la mayor representatividad en su conjunto, mientras que el muestreo por conveniencia es el que anuncia la mayor cantidad de sesgos en sus resultados.

La exactitud que se requiere en los resultados de un estudio está alineada con la representatividad de la técnica de selección, mientras que **la factibilidad** está alineada con la presencia de sesgos.

El muestreo por cuotas

Divide a la población en grupos y selecciona cuotas, dentro de cada grupo, si el tamaño de la cuota está en proporción al tamaño del grupo, se denomina proporcional, pero si el tamaño de la cuota es el mismo independientemente del tamaño del grupo se denomina equilibrado, y si además la selección de los elementos que conforman la cuota se realiza de forma aleatoria, se denomina **cuasi probabilístico**.

Muestreo según criterio

Selecciona elementos basándose en el conocimiento o criterio del investigador, es el más representativo aquel que se realiza basado en el análisis, con una proyección por parte del investigador, pero si esta proyección no existe, se realiza de manera teórica y si incluso la teoría no existe, se realizará de forma intencional o de juicio.

El muestreo en bola de nieve

Selecciona a unos pocos elementos, quienes recomiendan a otros y así sucesivamente, puede ser hasta completar una cuota y de forma dirigida o adaptativo y flexible mientras dure su ejecución o de manera simple, superficial y con incentivos.

El muestreo por conveniencia

Selecciona a los elementos más accesibles o fáciles de contactar, pudiendo ser de contexto u oportunidad, de voluntariado o incluso accidental

Veamos un ejemplo para cada uno de estos conjuntos:

1. Muestreo por cuotas

Vamos a suponer que, queremos realizar un estudio de encuesta en una universidad, en la que incluimos cinco escuelas profesionales y también 5 años de estudios, construyendo así una tabla de contingencia de 25 núcleos.

Es un muestreo de cuotas **equilibradas**, porque el tamaño de cada cuota es exactamente el mismo y es **no cruzadas** porque cada escuela profesional, así como cada año de estudio participan en una sola ocasión y será **cuasi probabilístico** si los elementos dentro de cada cuota se seleccionan al azar.

2. Muestreo según criterio

Vamos a suponer que estamos realizando un experimento, en el cual los participantes están divididos en cuatro grupos etarios, provenientes de cuatro localidades, construimos una tabla de contingencia de 16 núcleos.

Si cada tratamiento tiene cuatro réplicas necesitaríamos 64 unidades experimentales, pero si consideramos a cada tratamiento en un solo grupo etario y en una sola localidad, lo resumimos a 16, estamos hablando del **diseño cuadrado latino**, muestreo según criterio.

3. Muestreo en bola de nieve

Puede seleccionarse de **forma intencional** al inicio o arranque de cada bola de nieve o puede ser también de **forma accidental**, es lineal si cada participante recomienda únicamente a un elemento adicional.

Pero puede ser exponencial si cada participante recomienda a más de uno, es no discriminatorio si se incluyen a todos los recomendados, pero será discriminatorio si el investigador selecciona, a cuál de ellos incluir en la muestra y a cuál de ellos no.

4. Muestreo por conveniencia o por comodidad

Puede ser de contexto u oportunidad como encuestar a los trabajadores de la empresa el día del aniversario en que todos asisten.

Puede ser de voluntariado si invitas a los participantes de las redes sociales a completar un cuestionario, o errático o accidental si en este momento acudimos a un centro comercial para encuestar a los presentes.

Sesgos de selección

La validez de los resultados de un estudio, vienen siendo amenazadas por el error aleatorio y el error sistemático.

Cuando conseguimos reducir el **error aleatorio**, hemos alcanzado **mayor precisión**, y cuando conseguimos reducir el **error sistemático** hemos alcanzado **mayor exactitud**, para esto debemos controlar los sesgos de selección y sesgos de medición.

En esta presentación sesgos de selección...

En el orden en el que aparecen en el desarrollo de un estudio y son: sesgo de pertenencia, sesgo de admisión, sesgo de autoselección, sesgo de aleatorización y sesgo de prevalencia.

Vamos a suponer que queremos realizar un estudio de un programa educativo sobre los estudiantes universitarios, definimos como población estudiantes universitarios, si agregamos un criterio de inclusión que, indique estudiantes de universidades públicas estaríamos excluyendo a los estudiantes de universidades privadas cometiendo así el sesgo de pertenencia.

Incluso si los criterios de inclusión y exclusión son adecuados y seleccionamos estudiantes solamente de las universidades de la capital, estaríamos excluyendo a los estudiantes de universidades de provincias, cometiendo así el sesgo de admisión.

Sin perjuicio de lo anterior y encuestamos a los estudiantes universitarios únicamente voluntarios, estaríamos cometiendo el sesgo de autoselección

Pero si ya tenemos conformado, el grupo de los estudiantes que van a participar y manipulamos la asignación de los estudiantes a cada grupo, estamos cometiendo el sesgo de aleatorización

Finalmente, si analizamos los resultados solamente de los estudiantes que completan el programa educativo y no, de aquellos que lo abandonaron, estamos cometiendo el sesgo de prevalencia

1. Sesgo de pertenencia

Si la población de estudio está definida como mujeres embarazadas, y por el error colocamos el criterio de inclusión de mujeres embarazadas de color naranja, estaríamos excluyendo a las mujeres embarazadas de color verde y a las mujeres embarazadas de color celeste, aun cuando cumplen la definición de la población objeto de estudio, para prevenir ello, hay que redefinir los criterios de inclusión y exclusión.

2. Sesgo de admisión

Si la población de estudio está definida como varones de color naranja y en el estudio incluimos, únicamente a aquellos que se encuentren en nuestra localidad o en el área urbana, por nuestra comodidad.

Estaríamos cometiendo este sesgo, para mitigarlo debemos utilizar un muestreo estratificado o por conglomerados.

3. Sesgo de autoselección

Si la población de estudio está definida por varones de color naranja e incluimos en el estudio, únicamente a aquellos que resultan voluntarios.

Estaríamos excluyendo a aquellos que no decidieron participar en el estudio, sesgando así nuestros resultados. La única forma de eliminar este sesgo es utilizar un método de selección completamente aleatorio.

4. Sesgo de aleatorización

Si la población de estudio está conformada por varones de color naranja y varones de color celeste en la misma proporción, se entiende que, al construir el grupo experimental y el grupo control la proporción de varones de color naranja y color celeste debe ser la misma que en su población.

Si esto no ocurre, estamos cometiendo el sesgo de aleatorización y para su prevención debemos utilizar métodos de selección estrictamente aleatorios.

5. Sesgo de prevalencia

Si la población de estudio está constituida por ocho varones que se inscriben en un programa de dieta y ejercicios, cuyo objetivo es bajar de peso regentado por un gimnasio,

- En el primer mes, cuatro de ellos están obteniendo resultados tasa de éxito 50%.
- En el segundo mes uno de los que no está obteniendo resultado se retira tasa de éxito $4 \text{ de } 7 = 57\%$.
- El tercer mes otro de los que no está obteniendo resultados se retira tasa de éxito $4 \text{ de } 6 = 67\%$.

El gimnasio anunciará en su publicidad que su éxito en el programa es del 67% cuando en realidad, todos los meses se inscriben en el programa ocho personas, de las cuales únicamente cuatro consiguen sus resultados.

CAPÍTULO 8: LA RECOLECCIÓN DE DATOS

La documentación reúne información proveniente de fuentes escritas, digitales o institucionales que permiten contextualizar y sustentar el estudio. Incluye literatura científica, bases de datos, registros administrativos y otros documentos verificables. Su análisis crítico fortalece el fundamento conceptual y mejora la precisión interpretativa.

La observación consiste en registrar de manera sistemática el comportamiento, atributos o dinámicas del fenómeno en su propio entorno. Puede ser estructurada o no estructurada, participante o no participante, según el objetivo del estudio. Su valor radica en captar información directa, sin mediación declarativa del sujeto.

La entrevista permite obtener datos mediante una interacción comunicativa guiada por un propósito investigativo. Puede adoptar formatos estructurados, semiestructurados o abiertos, según el nivel de estandarización requerido. Su fortaleza es el acceso a percepciones, significados y experiencias que no emergen por simple observación.

La encuesta recoge información a través de cuestionarios aplicados a una muestra definida, con el fin de medir variables de interés. Requiere preguntas claras, indicadores bien operacionalizados y un procedimiento de aplicación estandarizado. Su potencia reside en la capacidad de generar datos comparables y analizables estadísticamente.

La documentación

En investigación, documentar es **obtener información de fuentes registradas previamente**, sin que el investigador haya participado en la medición original de los datos.

Consiste en **acceder a documentos** cuyo fin primario no fue investigativo, como clínicos o administrativos, y extraer de ellos los valores necesarios según los objetivos del estudio.

Se aplica, cuando los eventos ya ocurrieron o cuando por razones éticas o de factibilidad, no se puede repetir el evento o generar la situación observada.

1. La documentación suministra datos secundarios

Existen dos tipos de datos, **los primarios** que son registrados por el investigador y **los secundarios** que ya han sido con fines ajenos al estudio.

La documentación suministra datos secundarios porque el investigador no realiza nuevas mediciones, sino que accede a registros existentes con fines distintos al estudio en curso.

Al no participar en su medición, el investigador no puede garantizar a la precisión ni exactitud lo que define su condición de dato secundario.

2. La documentación no asegura la precisión y exactitud de los datos

Un dato preciso es aquel que al volverse a medir presenta poca variación. Esta estabilidad refleja el control del error aleatorio en el proceso de medición.

Un dato exacto es aquel que se aproxima al valor real. Se consigue controlando el error sistemático que surge por condiciones no estandarizadas.

La documentación no garantiza precisión, ni exactitud, ya que el investigador no participó en la medición original y por tanto, no controló dichos errores.

3. La documentación corresponde al estudio retrospectivo

En investigación los estudios pueden ser **prospectivos** o **retrospectivos**, según el control del investigador sobre la medición de la variable de estudio.

Los estudios prospectivos utilizan **datos primarios** recolectados por el investigador. En cambio, los estudios retrospectivos emplean datos secundarios ya registrados.

Cuando la **variable de estudio** proviene de datos secundarios, es decir, fue medida antes del inicio del estudio, este se clasifica como **retrospectivo**.

4. La documentación no necesita instrumentos de medición.

El estudio retrospectivo **no requiere instrumentos de medición**, ya que no genera nuevos datos, sino que se basa en información **previamente registrada** por terceros.

El investigador no realiza mediciones directas, sino que accede a valores ya consignados en fuentes como historias clínicas o bases de datos institucionales.

Dado que la medición ocurrió antes del estudio y sin su intervención, no necesita equipos como balanzas, termómetros o pruebas diagnósticas.

5. La documentación utiliza una ficha de recolección de datos

En los estudios retrospectivos, la ficha de recolección de datos se utiliza como una herramienta administrativa para organizar y sistematizar la extracción de datos secundarios.

Esta ficha no mide ni genera información nueva. Su función es trasladar los valores ya registrados desde la fuente original hacia una base de análisis.

Al no ser un instrumento de medición, **no requiere de validación**, ya que no interviene en el proceso de obtención, ni transformación del dato.

6. La documentación utiliza documentos físicos y digitales

La documentación tradicional recurre a **documentos físicos** como historias clínicas, informes o actas. Son útiles, aunque pueden ser incompletos y tener errores de registro.

Estos documentos, no fueron creados con fines investigativos, pero permiten desarrollar estudios preliminares como los exploratorios, descriptivos e incluso relacionales.

También se utilizan **documentos digitales** como bases de datos. Si provienen de mediciones directas, sus datos pueden considerarse.

7. Finalmente, no existe el estudio documental

La documentación es una técnica de recolección de datos, no un tipo de estudio. Por tanto, no existe el llamado estudio documental.

Un estudio puede emplear varias técnicas de recolección de datos como: la documentación, observación o encuesta. Ello no define su tipo ni su diseño.

La taxonomía de la investigación exige criterios **exhaustivos y excluyentes**. Clasificar estudios por técnica usada no cumple estas condiciones.

La observación

La observación es una técnica de recolección de datos que consiste en **registrar de forma intencionada hechos, conductas o rasgos** directamente observables. Puede ser estructurada o no estructurada y centrarse en variables objetivas, medibles o clasificables, o en categorías subjetivas predefinidas o emergentes.

Su valor radica en captar el fenómeno tal como ocurre, usando desde instrumentos formales hasta la percepción libre del observador.

En todo proceso de observación participan los siguientes elementos: el observador, el ente observado, los medios de observación y las circunstancias de la observación.

De acuerdo con estos elementos, podemos plantear las siguientes clasificaciones

- Según la relación entre el observador y el ente observado.
- Según los medios de observación.
- Según las circunstancias de la observación.

1. Según la relación entre el observador y el ente observado

La observación puede ser participante donde el investigador se involucra activamente con el contexto observado. Su presencia influye, pero permite mayor comprensión interna y no participante, donde el investigador se mantiene al margen del fenómeno.

Ejemplo de **observación participante**. El investigador se une como voluntario en un comedor popular. Se integra en las actividades cotidianas junto al resto del equipo local.

Durante su participación observa la organización de turnos, las decisiones internas y las interacciones entre usuarios sin interrumpir el flujo natural del lugar. Reconoce que su presencia influye en el entorno, pero valora el acceso directo a **significados desde dentro del fenómeno**.

Ejemplo de **observación no participante**. El investigador se ubica en una banca apartado del grupo. Observa a los niños jugar en la plaza sin acercarse ni involucrarse en sus actividades.

Permanece en silencio sin emitir juicio ni interactuar. Solo registra lo que ve, los juegos, las interacciones, las emociones que surgen espontáneamente. Todo lo anota en su cuaderno, buscando captar el comportamiento natural **sin alterar el entorno**. Su distancia garantiza objetividad y mínima influencia.

2. Según los medios de observación

La observación puede ser **sistemática** cuando los medios están definidos. Antes de iniciar se emplean instrumentos, materiales o fichas que permiten repetir y verificar lo observado o **asistemática** cuando no hay medios previamente establecidos. El investigador observa libremente sin reglas, confiando en su percepción directa.

Ejemplo de **observación sistemática**. El profesor observa y completa una ficha. Evalúa si el alumno parece motivado, distraído o desinteresado. Cada conducta está definida previamente en la guía de observación.

No improvisa, sino que aplica criterios establecidos para registrar lo que percibe. Aunque se trata de juicios subjetivos, el uso de categorías predefinidas y ficha estructurada permite sistematizar la observación con cierto control.

Ejemplo de **observación asistemática**. El entrenador observa el partido **sin ficha ni criterios previos**. Nota que un niño motiva y organiza espontáneamente a sus compañeros de equipo.

No esperaba registrar liderazgo, pero lo anota porque surge de manera natural. Su mirada se guía por lo emergente, no por una lista cerrada. Al no seguir un protocolo, capta detalles imprevistos. La flexibilidad de la observación asistemática permite descubrir comportamientos clave.

3. Según las circunstancias de la observación

La observación puede ser **controlada** cuando el investigador manipula las condiciones del entorno para observar con precisión o **no controlada** cuando el fenómeno se registra en su ambiente natural sin alterar nada.

Ejemplo de **observación controlada**. En un laboratorio se coloca a un niño solo frente a un dulce y se le indica que espere. El investigador observa desde otra sala sin intervenir.

Todo ha sido diseñado. Tiempo, espacio y condiciones. Se controla el ambiente para que el fenómeno ocurra de forma precisa. El registro se realiza con cámaras y cronómetro. La artificialidad del entorno permite medir con exactitud, aunque se pierda naturalidad.

Ejemplo de **observación no controlada**. Durante el recreo, un docente observa desde lejos cómo sus alumnos se agrupan espontáneamente. No interviene ni organiza la situación.

El docente solo registra lo que ocurre en el entorno natural. **sin alterar la dinámica social**. La escena no fue provocada. Lo observado es auténtico. El valor está en captar el comportamiento tal como ocurre.

La entrevista

La entrevista es una **técnica de recolección de datos** basada en la comunicación verbal entre el investigador y el entrevistado mediante un diálogo directo e intencionado.

Permite obtener **información primaria** sobre conocimientos, percepciones, experiencias o actitudes. Según su estructura, la entrevista puede ser: entrevista a profundidad, entrevista enfocada y entrevista estructurada.

1. La entrevista a profundidad

La entrevista a profundidad **es exploratoria**, parte de una sola pregunta y permite que el entrevistado desarrolle su discurso de forma libre y continua. Es netamente cualitativa y holística. Busca explorar y descubrir significados **sin seguir reglas rígidas** ni pretender tabular o clasificar datos. Su fin no es confirmar hipótesis, sino generar ideas que evidencien un concepto.

No tiene guion fijo. Las preguntas surgen a medida que emergen las respuestas. El investigador estimula el relato, sin imponer temas externos. Los límites los impone la propia investigación.

La entrevista se adapta al contexto sin restringirse a formato orden ni duración preestablecida. Requiere un entrevistador experto en el tema, capaz de adaptarse al sujeto y extraer lo relevante. La profundidad depende de su habilidad comunicativa.

Un ejemplo claro es la anamnesis, la entrevista inicial que el médico realiza a su paciente en la primera consulta para conocer sus principales molestias.

El médico inicia con una pregunta abierta como, ¿cuál es su principal malestar? A partir de ahí, explora síntomas y señales que orienten hacia un diagnóstico preliminar. Durante la anamnesis, el médico no sigue un guion rígido. Su indagación es flexible, guiada por juicio clínico.

2. La entrevista enfocada

La entrevista enfocada, **también llamada semiestructurada**, parte de una lista de palabras clave elaboradas previamente como guía para su desarrollo.

Dichas palabras clave pueden surgir de una entrevista a profundidad. No se permiten añadir nuevos temas fuera de los previamente establecidos en la planificación.

El objetivo es orientar la entrevista hacia aspectos específicos y evitar omisiones importantes, conservando el foco temático que la investigación requiere.

Tiene cierto grado de estructura, ya que sigue una lista de temas. Sin embargo, el entrevistador puede adaptar la forma y el orden de las preguntas. Aunque hay temas definidos, la entrevista permite libertad operativa para plantear las preguntas según el ritmo del diálogo y las respuestas del entrevistado.

El objetivo es cubrir todos los temas relevantes **sin forzar una secuencia fija**, permitiendo adaptación sin perder el foco investigativo central.

Un médico especialista realiza una entrevista enfocada limitándose a explorar solo lo correspondiente a su especialidad. **Busca plantear una hipótesis clínica, no confirmarla.**

La entrevista culmina con una impresión diagnóstica que requiere ser comprobada por otros medios. Lo importante no es el orden ni el contexto, sino si se logra o no formular la hipótesis. El objetivo es plantear la existencia del concepto en el sujeto.

3. La entrevista estructurada

La entrevista estructurada, también llamada **semicuantitativa**, busca demostrar la presencia de un concepto, no explora, confirma, descarta o clasifica.

El entrevistador no improvisa ni interpreta libremente. Sigue parámetros fijos que permiten evaluar hipótesis concretas. Solo se puede hacer preguntas previamente establecidas. Si el entrevistado responde, se permiten ampliaciones, pero no nuevos temas o desvíos.

Se parte de una guía con preguntas inmodificables. Lo que se pregunta está definido desde el diseño del estudio. Puede haber respuestas abiertas o cerradas, pero el control está en la formulación. Esto facilita su aplicación, comparación y análisis objetivo.

Se aplica cuando ya se tiene una hipótesis clara. La entrevista termina solo cuando se ha confirmado o descartado el concepto que se intenta validar.

En medicina, la entrevista estructurada **es como un algoritmo diagnóstico.** El médico psiquiatra formula preguntas para verificar signos y síntomas esperados. Cada síntoma forma parte de la hipótesis.

El especialista no puede omitir ninguno. El resultado debe ser diagnóstico **confirmado o descartado.** El resultado se convierte en un dato objetivo, sano o enfermo. Si se aplica a muchos pacientes puede generar análisis cuantitativos formales.

La encuesta

La encuesta usa un cuestionario con preguntas escritas, organizadas y estandarizadas que permiten recoger datos de forma sistemática y comparable.

Cada ítem debe ser claro y preciso para que todos los encuestados lo comprendan igual, **asegurando uniformidad** en la recolección de datos. Esta estructura permite codificar, tabular y analizar estadísticamente las respuestas, generando información cuantificable.

Su objetivo es captar respuestas frente a **preguntas diseñadas previamente**, reflejando conocimiento, opiniones, conductas o pensamientos medibles.

La encuesta mide frecuencia, intensidad o distribución de respuestas, por lo que se clasifica como **una técnica cuantitativa**.

Estudia conocimientos, actitudes o creencias como variables observables y al aplicarse a muestras representativas permite generalizar resultados.

1. Según la administración del instrumento puede ser: autoadministrada y heteroadministrada.

En la encuesta autoadministrada

El encuestado responde solo, sin guía. No hay encuestador que supervise, corrija o insista. Este tipo de encuesta otorga libertad, pero también genera riesgo. Puede fomentar respuestas sinceras o abandono total. Es silenciosa, eficiente e incierta.

Ejemplo, el profesor reparte un examen escrito y los alumnos lo responden sin ayuda. A su ritmo no hay intervención, solo vigilancia del tiempo límite.

En la encuesta hetero administrada

El encuestador no solo aplica el instrumento, también lo interpreta, lo lee y lo adapta al ritmo del encuestado en tiempo real. El control es su virtud y garantiza comprensión, pero también es su riesgo, pues cada palabra del encuestador puede influir en la respuesta final.

Ejemplo, el profesor lee en voz alta cada pregunta del examen y anota las respuestas dictadas por los alumnos. Él aplica, controla y media el instrumento.

2. Según el tiempo de aplicación puede ser: sincrónica y asincrónica.

En la encuesta sincrónica

Todos responden al mismo tiempo. El instrumento se con inicio y fin definidos. El tiempo es parte del diseño, controla condiciones y asegura respuestas completas. El reloj garantiza eficiencia, pero limita reflexión.

Ejemplo, el profesor entrega el examen a las 10:00 am y lo recoge a las 10:30. Nadie responde antes ni después, solo dentro del tiempo asignado.

En la encuesta asincrónica.

El tiempo no es compartido. Cada encuestado responde cuando puede, quiere o se acuerda. Aquí el reloj no coordina, solo observa. Se gana libertad, pero se pierde control. El investigador no sabe cuándo llegarán las respuestas ni si llegarán. El dato flota entre intención y abandono.

Ejemplo, el profesor envía un examen virtual con plazo de una semana. Algunos alumnos responden el mismo día, otros lo olvidan.

3. Según el modo de aplicación puede ser: presencial y no presencial.

En la encuesta presencial

El dato se obtiene frente a frente. El encuestador observa, guía y recoge no solo respuestas, sino también gestos, silencios y dudas.

Es costosa y lenta, pero profundamente precisa. Nada sustituye la mirada directa ni la presencia compartida.

Ejemplo, el profesor entrega un examen impreso en clase, observa cómo responden sus alumnos, aclara dudas en el acto y recoge los resultados al final.

En la encuesta no presencial

El instrumento viaja sin cuerpo, **no hay mirada**, solo una interfaz que espera ser leída. El dato llega sin contexto, pero llega.

El investigador no ve al encuestado ni sabe cuándo responde. Se gana alcance y rapidez, pero se pierde control. La distancia da libertad y también silencio.

Ejemplo, el profesor envía un examen por Google Forms. Los alumnos lo completan desde casa sin supervisión.

El examen que aplica un profesor a sus alumnos utiliza un cuestionario, aplicándolo como encuesta autoadministrada, sincrónica y presencial.

CAPÍTULO 9: LOS INSTRUMENTOS

Los instrumentos de medición constituyen las herramientas mediante las cuales se obtienen datos válidos y confiables sobre las variables del estudio. Pueden ser cuestionarios, escalas, inventarios o equipos especializados, según la naturaleza del fenómeno. Su calidad determina directamente la solidez de los resultados.

La creación de instrumentos requiere definir claramente el contenido que debe medirse y asegurar su congruencia con las dimensiones de la variable. Implica formular ítems pertinentes, representativos y comprensibles para la población objetivo. Este proceso garantiza que el instrumento refleje fielmente el constructo teórico.

Las propiedades métricas de los instrumentos evalúan su calidad psicométrica mediante criterios de validez y confiabilidad. La validez analiza si el instrumento mide lo que pretende medir; la confiabilidad, la estabilidad y consistencia de sus resultados. Juntas permiten determinar la precisión y utilidad del instrumento en investigación.

Los sesgos de medición surgen cuando errores sistemáticos alteran las puntuaciones obtenidas, desviando la medición del valor real. Pueden originarse en el instrumento, en el evaluador, en el participante o en el contexto de aplicación. Identificarlos y controlarlos es esencial para preservar la integridad de los datos recogidos.

Instrumentos de medición

Los instrumentos de medición miden, si no entregan un valor final de medición no son instrumentos.

Clásicamente tenemos a los instrumentos documentales e instrumentos mecánicos los **instrumentos documentales** evidencian dimensiones lógicas correspondientes a indicadores subjetivos, mientras que los **instrumentos mecánicos** evidencian magnitudes físicas correspondientes a indicadores objetivos.

En esta presentación instrumentos documentales...

Los instrumentos de medición documentales son: los cuestionarios las escalas y los inventarios.

- Los cuestionarios tienen un resultado de medición dicotómico como ingresar o no ingresar a la universidad.
- Las escalas tienen un resultado de medición ordinal como el dolor leve, moderado y severo.
- Los inventarios por su lado tienen un resultado de medición politómico como tener vocación para las ciencias de la salud las ciencias sociales o las ingenierías.

Los cuestionarios y las escalas son univectoriales de una sola dirección mientras que los inventarios son multivectoriales en varias direcciones

El cuestionario está conformado por un conjunto de preguntas que perciben evaluar alguna capacidad por ejemplo cognitiva. Un examen de ingreso a la universidad está conformado por un conjunto de preguntas, cada pregunta tiene un resultado dicotómico acierta y no acierta del mismo modo el resultado final también es dicotómico aprobado y desaprobado, ingresa o no ingresa a la universidad.

Las escalas están conformadas por un conjunto de reactivos sobre los cuales los evaluados indican su grado de acuerdo o desacuerdo. Como las escalas de actitudes frente al aborto terapéutico de acuerdo, indiferente, en desacuerdo, cuyo resultado ordinal también se expresa en el valor final de medición del instrumento.

El inventario está conformado por un conjunto de ítems clasificatorios donde ninguna respuesta es correcta o incorrecta como tu deporte favorito: fútbol, basquetbol, voleibol, natación, etcétera.

La creación y validación de instrumentos corresponde a una línea de investigación que atraviesa etapas, fases o momentos:

La primera fase es netamente cualitativa y se denomina validez de contenido, corresponde a la creación del instrumento.

La segunda fase es cuantitativa porque hace uso de las herramientas estadísticas para evaluar sus propiedades métricas del instrumento previamente creado.

La validez de contenido se refiere a la medida en que un instrumento de medición refleja adecuadamente el dominio del constructo que se pretende medir. Asegura que todas las dimensiones de un concepto estén representadas en el instrumento evitando contenido insuficiente o impertinente.

Vamos a suponer... que queremos evaluar el nivel de conocimientos sobre operaciones aritméticas básicas en un conjunto de niños. Las dimensiones son: suma, resta, multiplicación y división. Si omitimos la dimensión división estamos frente a un contenido insuficiente y por tanto **no válido**. Si agregamos la dimensión radicación estamos frente a un contenido no pertinente y por tanto **no válido**, es posible que el instrumento pueda tener contenidos insuficientes y no pertinentes agravando la situación.

- El ejemplo de las operaciones aritméticas corresponde a un escenario ideal donde el concepto está plenamente definido le vamos a llamar **escenario uno**.
- Es posible que el concepto que queremos evaluar esté parcialmente definido exista teoría no definitiva e incluso existan algunos instrumentos le llamaremos **escenario dos**.
- Finalmente es posible que el concepto aún no esté definido no exista teoría previa o muy escasa lo vamos a denominar **escenario tres**.

Vamos a utilizar la validez racional, la validación por jueces y la validez de respuesta para asegurar la validez de contenido.

Cuando el concepto está plenamente definido se elabora el instrumento en función a la teoría oficial y decimos que hemos alcanzado la validez de contenido a través de la validez racional.

Cuando el concepto está parcialmente definido se elabora un instrumento en función a la teoría disponible y se hace evaluar por un conjunto de jueces, hablamos de la validación por jueces.

Cuando el concepto no está definido se elabora un instrumento con la entrevista abierta a la población objetivo, para luego asegurar que estos contenidos los puedan verificar un conjunto de expertos.

Validación de instrumentos

Un diseño de validación de instrumentos atraviesa diferentes etapas o momentos. La primera fase es netamente cualitativa y se denomina **validez de contenido** en la cual se lleva a cabo la construcción del instrumento, al mismo tiempo, luego hay que evaluar sus propiedades métricas, para lo cual haremos uso de la estadística, constituyendo así, una segunda fase cuantitativa.

Validez de constructo

La validez de constructo evalúa si el instrumento refleja la **estructura teórica** del concepto si el constructo tiene varias dimensiones el análisis estadístico debe respaldar que los ítems se agrupan correctamente según esos dominios.

La estructura factorial se refiere a la relación que tienen los ítems entre sí y respecto de su propia dimensión, así como del instrumento propiamente dicho. La validez de constructo se alcanza de acuerdo al grado de conocimiento que se tiene acerca del concepto que se pretende medir

- Escenario uno: cuando **el concepto está plenamente** definido se conocen las dimensiones que conforman el concepto. Por tanto, la validez de constructo está asegurada.
- Escenario dos: cuando **el concepto está parcialmente definido** se plantean las dimensiones con las que se mide el concepto y se desarrolla el análisis factorial confirmatorio.
- Escenario tres: cuando **el concepto no está definido** no se pueden plantear dimensiones para medir el concepto y se desarrolla el análisis factorial exploratorio.

Fiabilidad o confiabilidad

Equivalente al término de la **fiabilidad** evalúa la **consistencia interna** a través del alfa de Cronbach y la estabilidad considerada como **consistencia externa**.

Consistencia interna evalúa en qué medida la variabilidad de los ítems influyen en la variabilidad total. Si en una matriz de datos calculamos la variabilidad para cada uno de los ítems, así como el alfa de Cronbach total y retiramos al ítem que expresa mayor variabilidad el **alfa de Cronbach se reduce**.

Mientras que si retiramos al ítem que expresa menor variabilidad el **alfa de Cronbach aumenta**. Esto porque el **Alpha de Cronbach** es una relación entre la sumatoria de la variabilidad de cada ítem respecto de la variabilidad total y se considera aceptable cuando está por encima de 0,8.

Estabilidad

Reúne los conceptos de **repetibilidad** estabilidad intra evaluador y **reproducibilidad** estabilidad entre evaluadores.

- **La repetibilidad** es el estudio de la variación de las mediciones obtenidas con un instrumento de medición utilizado varias veces por un operador, cuando se mide la misma característica en un mismo individuo.
- **La reproducibilidad** es el estudio de la variación de las mediciones de diferentes operadores que, utilizan un mismo instrumento de medición, cuando miden la misma característica en un mismo individuo.

Validez de criterio

Mide el grado en que los resultados obtenidos con un instrumento de medición se asocian o correlacionan con los resultados de una variable externa que se supone es la medida oficial referente válido o gold estándar.

Agrupamos la **validez concurrente** que evalúa la relación entre los resultados del instrumento y el criterio en el mismo momento y la **validez predictiva** evalúa que también el instrumento puede predecir un resultado futuro relacionado con el criterio.

Si la variable analizada o resultado de medición **es categórica** realizamos el índice kappa de Cohen, mientras que si la variable analizada o resultado de la medición **es una variable numérica** realizamos el índice de correlación r de Pearson

Rendimiento diagnóstico

Se refiere a la capacidad para identificar correctamente una condición o estado de interés en términos de precisión y exactitud, considera a los conceptos de sensibilidad y especificidad, para la construcción de una **curva de rendimiento diagnóstico** donde, utilizando el plano cartesiano colocamos a la sensibilidad en el eje de las ordenadas y a 1-especificidad en el eje de las abscisas.

Construimos así la curva de rendimiento diagnóstico con cada uno de estos pares ordenados, para luego proceder a calcular el área bajo la curva, mientras más alto sea mayor rendimiento diagnóstico.

Ubicamos el punto de la curva que sea más distante de la diagonal y este señala **el mejor punto de corte** en términos de sensibilidad y especificidad para el test.

Adaptación de instrumentos

- La adaptación de instrumentos consiste en modificar a las herramientas de medición, conservando su validez al aplicarlas en contextos diferentes al original.
- Para adaptarlos correctamente se requiere identificar qué aspectos podrían cambiar entre un contexto y otro y así adecuar cuidadosamente el instrumento.
- De esta forma se garantiza que las mediciones sean replicables en cualquier escenario donde se apliquen.

Si bien, los contextos a los cuales hay que adaptar los instrumentos son múltiples y además se superponen, los vamos a enlistar, según la frecuencia en la que se requieren: adaptación lingüística, cultural, sociodemográfica, temporal, normativa, teórica y tecnológica.

1. Adaptación al contexto lingüístico

Se refiere a: adecuar el instrumento **a otro idioma** asegurando que los términos y expresiones conserven el significado original.

En los años 90 Pepsi tradujo al mandarín su eslogan "Come alive with Pepsi" revive con Pepsi. Resultando en la frase Pepsi trae a tus ancestros de la tumba causando evidente confusión y rechazo en China por una mala traducción.

2. Adaptación al contexto cultural

Consiste en ajustar el contenido del instrumento para que sea **comprensible y pertinente en la cultura de destino** respetando los valores creencias y costumbres locales.

Un mexicano abrió en Perú un negocio, ofreciendo tortas mexicanas sus primeros clientes peruanos pidieron velas, dispuestos a cantar "cumpleaños feliz" creyendo que traerían un pastel, sin saber que en México la torta es un sándwich.

3. Adaptación al contexto sociodemográfico

Se enfoca en las características de la población objetivo como: edad, género, nivel educativo entorno socioeconómico, a fin de que resulte apropiado y accesible a ese perfil de personas.

Un regidor de un municipio periférico limeño, copió una ordenanza de la municipalidad de Surco, que obligaba a los vecinos a **pasear a sus perros con correa**, al intentar sustentarla el propio alcalde respondió al regidor "Aquí los perros se pasean solos".

4. Adaptación al contexto temporal

Significa actualizar el instrumento **según la época o momento vigente** sustituyendo referencias, ejemplos o datos desfasados para mantener su vigencia en el presente.

Durante un examen de estadística el profesor preguntó: Explique cómo se compara el z calculado con el z de la tabla. Los alumnos respondieron “nosotros usamos directamente el p -valor obtenido del software”.

5. Adaptación al contexto normativo

Implica **ajustar un instrumento para que cumpla con las leyes** normas o regulaciones del lugar, donde se aplicará para asegurar que, esté acorde con las exigencias legales vigentes.

Un japonés, llega a Perú con su licencia de conducir japonesa, timón a la derecha, pero al intentar validar su permiso, le indican que debe hacer otro examen adaptado a las reglas peruanas, timón a la izquierda

6. Adaptación al contexto teórico

Significa **alinear el instrumento con los cambios o avances** ocurridos en las teorías científicas que fundamentan el constructo evaluado garantizando así la vigencia conceptual

En agosto de 2006 un profesor de secundaria preparaba su habitual examen sobre astronomía “enumere los nueve planetas del sistema solar” cuando repentinamente se enteró de que la Unión Astronómica Internacional había decidido que Plutón ya no era planeta.

7. Adaptación al contexto tecnológico

Implica **modificar el instrumento para su uso en un medio o plataforma diferente** optimizando su formato e interacción por ejemplo de papel a digital sin alterar lo que mide.

En marzo de 2020, un psicólogo se disponía a evaluar el clima laboral de los trabajadores en formato impreso, con la llegada de la pandemia y el trabajo remoto tuvo que transformar su instrumento, en un formulario digital.

Cuando hablamos de adaptación de instrumentos **incluimos también a la calibración** correspondiendo: la adaptación a la adecuación cualitativa y la calibración a la revisión de sus propiedades métricas, para asegurar precisión y exactitud en las mediciones.

Sesgos de medición

Para obtener **mediciones replicables** es necesario asegurar el control de los sesgos de medición y si, en un paso previo hemos controlado los sesgos de selección habremos conseguido reducir el error sistemático y si, antes de eso hemos controlado el error aleatorio hablamos de **la validez del estudio**.

Todo proceso de observación contiene los siguientes elementos: el observador, los medios de observación, el ente observado y las circunstancias de la observación. Esto nos permite **clasificar a los sesgos de medición** más frecuentes de acuerdo al orden en el que aparecen como: los relacionados con los medios de observación con el observador y con el sujeto observado y estos son: sesgo de precisión, sesgo de exactitud, sesgo del observador, sesgo de memoria y sesgo de adherencia.

Imagina que quieres evaluar el peso de un conjunto de niños utilizando una balanza doméstica en lugar de una balanza clínica acabas de cometer el **sesgo de precisión**. Incluso si cuentas con una balanza clínica, pero esta está descalibrada aparece el **sesgo de exactitud**.

Incluso si tienes una balanza clínica calibrada pero inconscientemente modificas las mediciones para intentar probar tu hipótesis acabas de cometer el **sesgo del observador**. Las personas con alguna patología suelen recordar con más ahínco lo negativo, presentándose así el **sesgo de memoria**. Los pacientes más adherentes al tratamiento suelen mejorar más incluso con placebo, a esto se le conoce como el **sesgo de adherencia**

1. Sesgo de precisión

Conocido también como **el sesgo de la capacidad diagnóstica del instrumento** son errores sistemáticos debidos a la falta de sensibilidad y especificidad en un método diagnóstico llevando a una clasificación incorrecta de los sujetos como falsos negativos y falsos positivos.

Para prevenirlo se debe **utilizar el instrumento Gold Standard** o patrón debidamente validado con alta sensibilidad y especificidad.

2. Sesgo de exactitud

Conocido también como **sesgo del rendimiento del instrumento**, aunque el instrumento registre mediciones estables y repetibles, estas se encuentran siempre desviadas en la misma dirección ya sea por encima o por debajo del valor real que se pretende conocer.

Para prevenirlo es necesario **una calibración periódica** y mantenimiento preventivo de los instrumentos de medición.

3. Sesgo del observador

Llamado también **sesgo del investigador** y este ocurre cuando el investigador ya sea consciente o inconscientemente deja que sus expectativas influyan en la recolección e interpretación de datos generando resultados sesgados hacia sus propias hipótesis

Para prevenirlo hay que implementar **estudios doble ciegos** donde el evaluador desconozca los objetivos del estudio.

4. Sesgo de memoria

Conocido también como el **sesgo del recuerdo** los evaluados recuerdan eventos pasados de manera diferente según su estado actual de salud. Por ejemplo, las personas enfermas tienden a recordar con mayor detalle a ciertas exposiciones previas comparadas con las personas sanas.

Para prevenirlo es necesario **limitar el periodo de recuerdo** porque entre más corto sea el intervalo solicitado menor será el error de memoria.

5. Sesgo de adherencia

Conocido también como **sesgo de adaptación o cumplimiento** surge cuando los participantes que siguen fielmente un tratamiento o protocolo, presentan características diferentes de quienes no lo cumplen adecuadamente habitualmente ventajosas.

Para prevenir esta situación es necesario aplicar un seguimiento personalizado integrando métodos técnicos, educativos y conductuales

Dicho todo esto es necesario diferenciar la precisión y exactitud de las mediciones respecto de la precisión y exactitud de los resultados del estudio.

La precisión y exactitud de las mediciones permite conseguir mediciones replicables prerequisite para conseguir la precisión exactitud de los resultados del estudio

En síntesis...

- Hay una precisión y exactitud **para las mediciones** y...
- Hay una precisión y exactitud **para la estimación de parámetros** a partir de los resultados del estudio.

CAPÍTULO 10: LOS RESULTADOS

La presentación de resultados organiza los hallazgos de manera clara, ordenada y fiel a los datos obtenidos, utilizando tablas, gráficas y descripciones sintéticas. Su función es exponer la evidencia sin interpretaciones ni juicios adicionales. La estructura debe corresponder al diseño, variables y objetivos planteados.

La interpretación de resultados explica el significado de los hallazgos en relación con las variables, el nivel investigativo y la lógica del estudio. Implica analizar patrones, contrastes y magnitudes para comprender su relevancia empírica. Este paso convierte los datos en información comprensible sin vincularla al propósito del estudio.

La discusión de resultados contrasta los hallazgos con la literatura previa y el fundamento teórico, evaluando coincidencias, discrepancias y aportes. Además, reconoce limitaciones metodológicas que influyen en la fuerza de la evidencia. Su objetivo es situar el estudio dentro del conocimiento existente y justificar su contribución.

Las conclusiones y recomendaciones sintetizan los aportes centrales del estudio y proponen acciones futuras basadas en los hallazgos. Las conclusiones responden directamente a los objetivos; las recomendaciones orientan aplicaciones prácticas o nuevas investigaciones. Ambas cierran el ciclo metodológico con claridad y pertinencia.

Presentación de resultados

- Los resultados, son el núcleo del trabajo de investigación.
- Si no se presentan resultados, no hay nada que comunicar.
- Sin resultados no existe informe final de investigación.

El principio rector, es la **comunicación con precisión**. El lector debe comprender los resultados, sin necesidad de recurrir a, otras secciones del reporte. La mejor forma es la más simple, ya sea texto, tablas o gráficas.

Presentación de resultados en texto

Texto como resultado. Responde a los objetivos planteados. El texto debe seguir una estructura lógica con contenido suficiente, pertinente, claro y preciso. En algunos casos es posible utilizar solo texto como forma de presentar los resultados de manera completa, sin depender de apoyos visuales.

- **Texto para estudios cualitativos.** Cuando no hay procedimientos estadísticos, el texto organiza los hallazgos observados. El texto consigue mostrar evidencia y responder a los objetivos.
- **Texto para datos simples.** Los datos simples como frecuencias, porcentajes y promedios se presentan en texto cuando son pocos, en este caso no se requiere de una tabla.
- **Texto que amplía la tabla o gráfica.** Cuando el uso de apoyos visuales es necesario, el texto no los repite, sino que destaca los elementos clave o añade la información que no está contenida en ellos.
- **Texto que presenta el análisis estadístico.** En estudios cuantitativos, el texto informa el valor del estadístico y el p-valor o lo que corresponda cuando no están incluidos en la tabla o gráfico.

Presentación de resultados en tablas

Las tablas permiten presentar gran cantidad de datos de forma resumida, ordenada y visualmente clara. Son clave cuando el texto no puede comunicar con eficacia la totalidad de los datos sin volverse extenso y/o confuso.

Una buena tabla debe ser autosuficiente. Debe incluir un título claro, categorías bien definidas y unidades de medida. Se puede complementar con texto la información que no se detalla en la tabla.

- **En los estudios descriptivos,** las tablas presentan frecuencias, porcentajes, promedios y desviaciones estándar. son útiles para caracterizar un conjunto de unidades de estudio. Ejemplo, una tabla con distribución por sexo, edad y nivel educativo, permite visualizar de forma rápida cómo está conformada la población analizada.

- **En los estudios analíticos**, las tablas evidencian relación entre variables ratificado por los procedimientos estadísticos como chi cuadrado o t de Student, incluyendo estimadores o el valor de p-valor según corresponda. Ejemplo, una tabla de contingencia que presenta el tabaquismo y la hipertensión y reporta valores de chi cuadrado y p-valor, permite comprobar la significancia estadística.

Presentación de resultados en gráficas

Las gráficas son herramientas visuales, que permiten comunicar resultados de forma muy rápida e intuitiva. Las gráficas deben contar con un título descriptivo, ejes claramente rotulados y un uso adecuado del color según su función. La información no representada en la gráfica puede complementarse mediante texto.

- **En los estudios descriptivos**, las gráficas son útiles para mostrar frecuencias o distribuciones de frecuencia, comparaciones entre grupos o proporciones destacadas de una variable. Ejemplo, un gráfico de barras que muestre el nivel de actividad física por grupos de edad, permite ver fácilmente qué grupo es el más sedentario.
- **En los estudios analíticos**, ayudan a visualizar la relación entre variables como las asociaciones, correlaciones o efectos de una variable sobre otra, con diagramas de dispersión o líneas de regresión. Ejemplo, un diagrama de dispersión entre el estrés y horas de trabajo con línea de tendencia positiva permite ver que a mayor carga laboral aumenta el nivel de estrés.

Presentación del análisis estadístico

Presenta solo los resultados finales. No se detallan los procedimientos estadísticos, solo se presentan los productos finales. Se presentan estadísticos como estimadores y p-valores. Evita mostrar salidas completas del software, no incluye todos los resultados del software estadístico. Se presenta únicamente lo esencial para responder al objetivo.

Escribe el valor exacto de p. El p-valor indica la probabilidad de cometer el error tipo I. No basta con decir p-valor menor a 0,05. Debe reportarse el valor exacto de p. Presenta las estimaciones con intervalos de confianza. En los estudios diferenciales reporta la media o proporción junto con su intervalo de confianza al 95% que en estadística representa la precisión.

Procedimientos estadísticos en anexos. Los procedimientos y la demostración de supuestos no se incluyen en la sección de resultados. Si se desea documentarlos deben ir en anexos. Ejemplo, una tabla reporta $X^2 = 10,74$, p-valor = 0,0047, lo cual demuestra que existe una asociación significativa entre el grado de instrucción y la automedicación según su nomenclatura.

Interpretación de resultados

La interpretación de resultados parte de una **descripción** reflexiva, no repetitiva, que destaca lo relevante y prepara el terreno para el análisis estadístico.

La interpretación del **análisis** estadístico actúa como puente hacia las conclusiones, articulando los resultados con argumentos que anticipan la discusión final.

La **interpretación** estadística es una conclusión intermedia basada en probabilidades que aún no es definitiva, pero orienta y enriquece la discusión posterior.

1. Descripción de resultados

La descripción de resultados no consiste en repetir cifras ni transcribir tablas, sino remarcar hallazgos clave con base en el discernimiento.

Es todo un proceso creativo. No basta dominar el método o la estadística. **Se requiere experiencia en la línea de investigación.** Solo así se elige con precisión qué resaltar y qué omitir.

El exceso de descripción, no garantiza calidad. Como el núcleo del estudio está en los resultados, una buena descripción los convierte en argumento científico.

En los estudios no inferenciales como los exploratorios o descriptivos de caracterización, la descripción de resultados es el insumo principal para interpretar.

En estos estudios no se aplican pruebas estadísticas, por lo tanto, la interpretación se construye directamente desde la lectura argumentada de los datos descritos. Así, al describir o interpretar sin inferencia, el investigador extrae conclusiones legítimas sustentadas únicamente en la **observación estructurada**.

En los estudios inferenciales, la descripción de resultados actúa como antesala al análisis estadístico, destacando diferencias o patrones que justifican aplicar dichos procedimientos. **no se limita a leer datos**, sino que selecciona lo relevante para preparar al lector al proceso inferencial como en chi cuadrado o regresión lineal.

Este puente analítico anticipa la dirección del hallazgo, permitiendo prever conclusiones que luego serán confirmadas con la probabilidad y la estadística.

2. Análisis estadístico

En el estudio descriptivo, el análisis estadístico permite extraer conclusiones simples, pero claves, ¿Existen diferencias o no según el objetivo del estudio? En todos los casos no es necesario mostrar el proceso completo del contraste de hipótesis. Basta con señalar el procedimiento usado, el estadístico calculado y el p-valor.

En el estudio relacional, el análisis estadístico identifica asociaciones entre variables. Una de las pruebas más comunes es la correlación de Pearson. En este caso, no se debe interpretar un coeficiente r por sí solo. Un valor alto de r no implica Relación, si el p-valor no indica significancia estadística. De modo que, primero se analiza el p-valor, en caso de que exista significancia estadística, recién se interpreta la magnitud y dirección de la correlación.

En el estudio explicativo, el análisis estadístico busca demostrar relaciones causales. Compara grupos aleatorizados antes y después de una intervención. Si tras la intervención se observan diferencias estadísticamente significativas y la estadística descarta relaciones espurias, las diferencias observadas se atribuyen a la manipulación. **La estadística no prueba causalidad por sí sola.**

3. Interpretación estadística.

La interpretación estadística **responde a hipótesis técnicas**, no al propósito sustantivo del estudio. Su lenguaje es probabilístico, no explicativo ni causal. Un p-valor significativo indica que un resultado no es producto del azar, pero no explica por qué ocurre ni qué implicancias tiene fuera del modelo estadístico. Confundir interpretación estadística con interpretación teórica lleva a errores. La estadística da señales, pero el sentido lo otorga el investigador.

El análisis estadístico exige un orden lógico. Primero se verifica si hay significancia. Solo si el p-valor es menor al umbral se interpreta el resultado. No se deben interpretar valores aislados. Un coeficiente alto sin significancia no tiene validez. El sentido analítico surge de la coherencia técnica. Solo se analizan los indicadores vinculados al objetivo del estudio. **La estadística no es acumulativa**. Más cálculos no implican mejor interpretación.

El análisis estadístico provee insumos, pero **no responde al propósito del estudio**. Esa tarea le corresponde al investigador que conoce la intención del estudio. Un analista de datos puede concluir que existe asociación, pero solo el autor del estudio puede afirmar si eso implica causalidad, utilidad o implicancia clínica. Interpretar con base en el propósito requiere integrar el resultado estadístico con el contexto, los objetivos y la lógica del planteamiento investigativo.

Discusión de resultados

La discusión de resultados es un proceso creativo en la que, el investigador integra: **evidencia empírica, marcos teóricos y razonamiento lógico** para interpretar el propósito del estudio, delimitar sus alcances, comparar los hallazgos con investigaciones previas y proponer aportes a la línea de investigación.

1. Interpretación del propósito

La valoración sustantiva de los hallazgos va **más allá del p-valor**. Busca interpretar los resultados a la luz del propósito que dio origen al trabajo de investigación.

Se empieza revisando **si los resultados responden a los objetivos** y si los datos confirman o rechazan la hipótesis planteada por el investigador. Los resultados esperados se analizan según el contexto teórico. Los inesperados deben ser explicados con cuidado, usando lógica y análisis crítico.

En cualquier caso, el investigador debe analizar qué significa el hallazgo y qué dice en verdad sobre el fenómeno que motivó el trabajo de investigación.

También hay que **analizar si el resultado apoya, cambia o amplía** las teorías que ya existen en el campo del conocimiento donde se desarrolló el estudio. Finalmente, es importante ver **cómo aplicar ese conocimiento** para resolver problemas reales o mejorar procesos de intervención y prácticas en distintos contextos.

2. Alcances y limitaciones

El alcance metodológico, muestra hasta dónde llegan los hallazgos según el diseño, la muestra usada y el contexto en el que se aplicó el estudio.

Los alcances precisan a qué población, tiempo y condiciones se aplican los resultados, ayudando a entender en qué casos son útiles o válidos. Reconocer **las limitaciones** da transparencia y fortalece la confianza en el estudio. Si se ocultan, se crean falsas certezas que debilitan el informe.

Es clave resaltar las fortalezas. Buen uso de instrumentos, control de sesgos y rigor en la recolección de datos fortalecen la confianza en los hallazgos. **Generalizar solo es válido si el diseño y la muestra lo permiten.**

Caso contrario, se debe aclarar que los resultados son válidos solo para la muestra estudiada en su propio contexto. En los estudios cualitativos no se busca generalizar, sino ver si los hallazgos pueden aplicarse en contextos parecidos. A eso se le llama **transferibilidad**.

3. Comparación con otros estudios

Comparar los resultados con estudios previos enriquece el análisis y permite ubicar el estudio dentro del debate actual en su campo científico. Aunque la comparación es cualitativa, hay que analizar críticamente la literatura, ver en qué se coincide o difiere y desde qué teoría se interpreta.

Si los resultados coinciden con otros estudios, se fortalece su validez. Si no hay que explicar con argumentos por qué ocurrió esa diferencia.

Las diferencias con otros estudios pueden deberse al método, la muestra o el contexto. Entenderlas bien requiere un análisis crítico y cuidadoso.

El investigador cita solo los estudios más relevantes para sus objetivos y diseño, priorizando la calidad antes que la cantidad de referencias. El contraste de los resultados con otros estudios es prerrequisito para armar un argumento sólido, que haga que el estudio aporte al debate científico con sentido y claridad.

4. Aportes a la línea de investigación

Proyectar el estudio al conocimiento exige ver con claridad, si aporta algo teórico, práctico o metodológico que ayude a avanzar en la línea de investigación. No todo resultado es un gran aporte.

Hay que analizar, si el estudio **explica, predice o aplica** de forma clara y coherente con su propósito. Aunque la muestra sea pequeña, un hallazgo positivo puede ser útil cuando confirma una hipótesis o permite plantear nuevas preguntas para investigar.

Desde este punto de la discusión surgen ideas para futuros estudios basadas en vacíos, límites o resultados inesperados que dejó la investigación actual.

Pero sugerir nuevas investigaciones no es una rutina.

Debe tener sentido y explicar cómo ayudaría a mejorar o contrastar lo que ya se encontró. De esta forma, el estudio en curso no queda aislado, sino que se conecta con otros estudios y aporta a su línea de investigación, otorgando la **relevancia científica**.

¡Atención al remate! Mientras que la interpretación de resultados con su análisis estadístico da respuesta a los objetivos del estudio del planteamiento operacional, solamente la discusión de resultados, con su interpretación al propósito, completan la intención del investigador expresada en el planteamiento del estudio.

ACERCA DEL AUTOR

José Supo es médico con formación especializada en bioestadística, metodología de la investigación y salud pública, con grado doctoral y reconocimiento como Doctor Honoris Causa. Ha desarrollado una sostenida labor académica como profesor invitado y conferencista en universidades de Perú, México, Colombia, Ecuador, Guatemala, diversos países de América Latina y Estados Unidos, consolidando una trayectoria internacional en docencia y capacitación científica.

Es autor de más de veinticuatro libros sobre investigación científica y análisis de datos, entre ellos obras de amplia difusión y uso académico en distintas ediciones. Asimismo, ha publicado el artículo “Taxonomía de la investigación en Salud Pública basada en la complejidad del análisis de datos” en la Revista Española de Salud Pública, en enero de 2026, contribuyendo al debate metodológico contemporáneo en el campo de la salud pública.

Su producción intelectual se complementa con programas formativos dirigidos a estudiantes y profesionales, orientados al fortalecimiento de competencias metodológicas y analíticas. Su trabajo se enfoca en elevar la calidad de la investigación científica en el ámbito hispanohablante, integrando rigor estadístico, claridad conceptual y aplicación práctica del conocimiento, con especial énfasis en la salud pública y las ciencias sociales.

DESCARGA ESTE MANUAL EN PDF

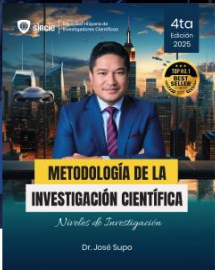
<https://bioestadistico.com/materiales>

JOSÉ SUPO



Médico, Docente e Investigador

OTRAS PUBLICACIONES



REDES SOCIALES

youtube.com/@bioestadistico

facebook.com/bioestadistico

tiktok.com/@bioestadistico

Más sobre el autor visita:

www.josesupo.com

ISBN: 978-612-99163-1-6



9 786129 916316



sincie

Sociedad Hispana de
Investigadores Científicos