TEMA 26

LA INVESTIGACIÓN CIENTÍFICA: CONCEPTO, LOS PRINCIPIOS DE CAUSALIDAD. DISEÑO DE UN PROYECTO DE INVESTIGACIÓN. EL MÉTODO ESTADÍSTICO: CONCEPTO. CLASIFICACIÓN DE LAS VARIABLES. ESTADÍSTICA DESCRIPTIVA. ESTADÍSTICA INFERENCIAL. TÉCNICA DE MUESTREO. ESTADÍSTICOS PARAMÉTRICOS Y NO PARAMÉTRICOS. LA INVESTIGACIÓN CUALITATIVA Y SUS MÉTODOS

INTRODUCCIÓN

La vida es un misterio, siempre ha sido así y siempre será así. Los seres humanos, y antes los primeros homínidos, hemos tratado de buscar el sentido de la vida a lo largo de los siglos en cada uno de sus ámbitos. Se han ideado todo tipo de estrategias que dieran sentido y explicación a aquello que vivíamos en el día a día, a lo desconocido.

La ciencia es un estilo concreto de pensamiento y acción. La investigación científica parte del supuesto de que el conocimiento disponible sobre un determinado tema o problema no es suficiente. Es necesario encontrar respuestas ciertas a cosas desconocidas y para ello se ha diseñado un método.

La ciencia no actúa sola. Se va generando pensamiento científico progresivo basado en el conocimiento anterior, en la generación de respuestas a lo desconocido. La ciencia avanza, y eso quiere decir que, la respuesta que se da en el siglo XIX a un determinado problema no es la misma que se ofrece en el siglo XXI. Y eso es porque la ciencia se sostiene en la cantidad de conocimiento que existe en cada momento. Cuanto más espacio iluminado sobre un determinado tema más posibilidades de hallar la respuesta correcta. A mayor conocimiento previo más posibilidades de obtener la respuesta correcta, a menor conocimiento previo, más probabilidades de fallo en la respuesta.

La ciencia se genera mediante equipos de investigación y conocimiento compartido, se basa en una red de discernimiento generado, debatido, discrepado, matizado, corregido, …

La ciencia no es estática, sino que siempre evoluciona y avanza. La ciencia no es infalible, necesita de una continua revisión. Los equipos de investigación deben estar dispuestos a ser críticos con sus descubrimientos y partir de la base de que, en el mejor de los casos, la respuesta que dan es cierta en este momento con un determinado conocimiento de la materia. Y siempre en un entorno cambiante que al ir variando necesita actualizar el conocimiento. Por eso el conocimiento científico puede quedarse obsoleto con facilidad y se caracteriza por el mismo movimiento que la misma vida que investiga.

1. LA INVESTIGACIÓN CIENTÍFICA: CONCEPTO

La investigación científica es un proceso ordenado y sistemático, de análisis y estudio. Utiliza diversos métodos o procedimientos específicos para el tipo de problema que se necesita resolver y para cada una de las partes del proceso de investigación.

A pesar de que es muy importante utilizar procedimientos concretos para cada problema las personas que investigan siguen una estructura general. El siguiente esquema reflejaría el proceso básico:

1. ¿Qué se sabe? Se reúne el conocimiento existente en relación con el tema que se va a investigar y, a partir de este conocimiento previo, se plantea el problema concreto que se va a intentar resolver.
2. ¿Qué se quiere saber? El problema necesita formularse como una pregunta precisa, concisa y ajustada. En el campo de la psicosociología se suele trabajar directamente con preguntas. En cambio, en el ámbito de las ciencias de la salud, a partir de la pregunta de investigación, se suelen formular una o varias posibles soluciones al problema que se denominan hipótesis. Las hipótesis deben estar fundadas en el conocimiento previo y ser precisas. Las hipótesis no son lo mismo que las suposiciones.
3. Se recopila la información a través de trabajo de campo, y en el caso de que se hayan formulado hipótesis, se someten a un contraste exhaustivo. Para que las consecuencias de las hipótesis sean contrastables se realizan mediciones, pruebas, encuestas, etc., dirigidas a comprobar su validez.
4. En el caso de que se estén utilizando hipótesis, se analizan y se hace una estimación y, en el mejor de los casos, se pueden considerar parcialmente válidas o verdaderas. Para ello los resultados obtenidos se someten a las pruebas estadísticas necesarias para aceptar o rechazar las hipótesis formuladas. Es necesario seguir contrastándolas en diferentes muestras o ensayos. En el caso de que no se formulen hipótesis la información se organiza, categoriza, codifica y triangula.
5. Las conclusiones obtenidas aumentan el conocimiento sobre el tema de estudio, generando un nuevo cuerpo de conocimiento, lo que permite que posteriores investigaciones planteen nuevas preguntas o problemas. Este nuevo punto de partida más evolucionado permite a la ciencia avanzar y seguir aumentando la sabiduría en el ámbito estudiado.

Dentro de la investigación científica, entendida como esa búsqueda intencionada de conocimientos o soluciones a los problemas planteados, nombramos el método científico utilizado especialmente en investigación experimental de ciencias de la salud.

La investigación científica puede perseguir un objetivo descriptivo de la realidad o puede desarrollar alguna teoría o regla. Ésta debe estar fundada e integrada en el sistema, siendo coherente con el resto de las reglas ya conocidas.

Hay ciertas características de la investigación científica que es necesario conocer a la hora de generar conocimiento:

- Debe realizarse mediante una metodología concreta aceptada por la comunidad científica que pueda considerar la validez de sus conclusiones, de manera que se pueda replicar, contrastar y refutar.
- Trata de reducir la subjetividad de las personas que investigan al mínimo. El objetivo es evitar los sesgos en la investigación al máximo posible de manera que, siguiendo el mismo procedimiento de investigación y con las mismas condiciones de medición o ensayo, se puedan llegar a las mismas conclusiones.
- El objetivo de la ciencia es ampliar el conocimiento existente, basándose en el pasado para aprender de él y generar nuevo. Es innovadora.

- El conocimiento permite elaborar teorías nuevas o describir y comprender fenómenos.

2. LOS PRINCIPIOS DE CAUSALIDAD

El principio de causalidad es un principio clásico de la filosofía y la ciencia, que afirma que "todo evento tiene, mínimo, una causa y un efecto, y que cualquier efecto que se produzca en la naturaleza debe estar precedido por, al menos, una causa".

Las cosas no ocurren de manera aislada, sino que unas están ligadas a otras en un proceso de interacción. Son interdependientes. Unas suceden a otras, y con frecuencia en el mismo orden. A los primeros sucesos en una relación los llamamos causas, y a los segundos efectos.

El principio de causalidad es un principio fundamental de la investigación científica, admitiendo que la mejor forma de entender y explicar los hechos es conocer sus causas, porque por un lado podemos prevenir las situaciones o los problemas y por otro controlar sus efectos.

Es importante volver a remarcar que los problemas, las situaciones a investigar, los hechos no suelen ser unicausales, sino que suelen producirse respondiendo a una multitud de causas. Algunas de ellas podremos verlas y estudiarlas por el conocimiento de base del que partimos en el momento de la investigación y otras permanecerán en el espacio sombrío de no conocimiento y aparecerán a medida que se vaya generando más conocimiento.

En ocasiones, tras la investigación, no se puede demostrar la causalidad entre dos variables o factores, pero sí se observa que existe un patrón que hace que esas dos variables se muevan conjuntamente. Una de ellas no causa la otra. En ese caso se puede valorar la correlación entre ellas.

3. DISEÑO DE UN PROYECTO DE INVESTIGACIÓN

Un proyecto de investigación es un proceso sistemático, organizado y objetivo, cuyo propósito es responder a una pregunta o hipótesis y así aumentar el conocimiento y la información sobre algo desconocido. Por lo tanto, un proyecto de investigación parte de la necesidad de dar solución a un problema con respuesta desconocida planteando una o varias preguntas o hipótesis que busquen aportar solucionar el problema.

Diseñar un proyecto de investigación es un proceso fundamental e imprescindible a la hora de enfrentar la búsqueda de respuestas. El proyecto es la guía o descripción del procedimiento de investigación que el equipo va a seguir.

Hay cuatro características claves hacia las que el equipo debe tender a la hora de plantear el diseño de una investigación:

1. La **neutralidad**. El diseño del estudio debe estar trazado para que los sesgos disminuyan lo máximo posible y aumente la objetividad.
2. La **fiabilidad**. Este concepto sirve, junto con el de validez, para evaluar la calidad de una investigación. La fiabilidad se refiere a la constancia en los resultados de una investigación. Cuanto más constante sea en la obtención de los mismos resultados utilizando los mismos métodos o instrumentos de medida en las mismas circunstancias más fiable será.

3. La **validez**. Se trata de la precisión con la que un método mide lo que tiene que medir. Una investigación con alta validez genera resultados que se corresponden con las propiedades, características y valores reales de lo medido.
4. La **generalización**. El diseño debe permitir la extensión de los resultados y conclusiones de la investigación de la muestra particular estudiada a la población de referencia.

Las etapas básicas en un proyecto de investigación son:

1. La elección de la idea o el problema de investigación.
2. El planteamiento del problema a través de la realización de un mapa de diseño de la investigación que establezca las líneas de trabajo, fije los objetivos, defina las preguntas de investigación, genere hipótesis de investigación (cuando sean necesarias) y explique la justificación del problema.
3. La fundamentación de un marco teórico y conceptual delimitando conceptos con su definición, haciendo una revisión de la literatura analizando el estado del problema, los antecedentes de la investigación y la identificación de lo conocido hasta el momento.
4. La definición de la metodología. A grandes rasgos se puede optar por una metodología cuantitativa, cualitativa o mixta. En cualquier caso, la metodología elegida debe justificarse y responder a los objetivos que persigue la investigación. En esta etapa también se define el alcance, las poblaciones o grupos a estudiar, el tamaño muestral, los métodos de medida o toma de datos e información, etc.
5. El diseño de un plan de trabajo. El cronograma es una herramienta indispensable a la hora de enfrentar una investigación con las actividades a realizar para cubrir los objetivos y trabajar de forma ordenada.
6. El desarrollo del proceso de investigación propiamente dicho donde se ejecuta todo lo proyectado en las etapas anteriores siguiendo el plan establecido y justificando las variaciones posteriores al plan inicial. En el proceso de investigación se obtendrá, se analizará y se describirá la información, para finalmente extraer conclusiones.
7. La exposición o publicación de la investigación. Este punto es vital para que la ciencia siga avanzando. Ya se haya ratificado la hipótesis de partida o se haya rechazado es muy importante que la comunidad científica sepa que así ha sido. Esta praxis contribuye al avance de la ciencia y de la sociedad.

La investigación se puede abordar, fundamentalmente, de tres maneras:

1. La **investigación cuantitativa** que mide la información y los datos, comprueba patrones, entiende si hay relaciones de causa-efecto o correlaciones, prueba hipótesis y, para todo ello, utiliza técnicas estadísticas. Los resultados se expresan numéricamente. Algunas de las herramientas propias de este tipo de investigación son: los experimentos, la observación, las encuestas, los cuestionarios o los registros.
2. La **investigación cualitativa** que recopila y analiza datos no numéricos para comprender conceptos, opiniones, experiencias y se ocupa de conocer los significados de una realidad. Los resultados se expresan en palabras. Algunas de las herramientas propias de este tipo de investigación son: las entrevistas, la observación, el análisis de documentos, videos, audios o situaciones presenciales, etc.
3. La **investigación mixta** combina ambas metodologías para conocer y comprender el problema. La ventaja fundamental es que la combinación de ambas ayuda a superar las

desventajas de cada método y ayuda a profundizar y obtener una imagen más completa de la situación de estudio.

4. EL MÉTODO ESTADÍSTICO: CONCEPTO

Un paso básico de cualquier investigación es la recogida y análisis de la información con el fin de o comprobar la hipótesis de estudio o describir y explicar un fenómeno. El método estadístico es la secuencia de procedimientos para el manejo de los datos de la investigación, tanto cualitativos como cuantitativos. Hay varias etapas dentro del método estadístico que se abordan de forma secuencial y dependen del diseño de la investigación. Las fases son: la recolección de los datos, su recuento o cómputo, la presentación de estos, la síntesis de la información y el análisis.

Primera etapa, **la recolección de datos**. El equipo investigador busca recoger información. La información se recoge que puede ser cuantitativa (datos numéricos como la talla, el peso, la edad, la concentración de una sustancia, etc.) o cualitativa (características descritas numéricamente como el género, la ocupación, la actividad económica, la opinión sobre un tema concreto, etc.). La información se recoge de diversas maneras en función del diseño del estudio: a través de la observación, mediante encuestas, con entrevistas individuales o grupales, mediante censos, etc. La calidad en esta fase es fundamental para disponer de datos exactos y confiables que fundamenten la investigación.

En este punto destacamos la *encuesta*, por ser ampliamente utilizada en el INSST. Se trata de un método de investigación estandarizado en el que se recopilan datos a través de un cuestionario prediseñado. Es una forma rápida que permite obtener y elaborar datos de modo rápido y eficaz. Es probablemente el método de investigación sociológico más utilizado, y es ampliamente usado también en la investigación médica.

Es importante distinguir la encuesta del cuestionario. Este último es la herramienta que usa la encuesta, pero la metodología de encuesta implica muchas etapas anteriores y posteriores a la aplicación del cuestionario, como la identificación del problema, la determinación del diseño de investigación, la especificación de hipótesis, la definición de variables, la selección de la muestra, el diseño del cuestionario, la organización del trabajo de campo, la obtención y tratamiento de los datos y el análisis de los datos e interpretación de los resultados.

Segunda etapa, **el recuento de datos**. Se cuantifica la frecuencia con la que aparecen las características de la investigación que son de interés para el propósito de ésta. Habitualmente se utilizan métodos informáticos para llevar a cabo el recuento de datos.

Tercera etapa, **la presentación de los datos**. Es necesario mostrar la información de forma sencilla de manera que sea clara y comprensible para cualquier persona ajena a la investigación. Las herramientas habituales de exposición de los datos son los cuadros o tablas y los gráficos.

Cuarta etapa, **la síntesis**. La información se resume para expresar de forma esencial los datos. En esta fase es importante extraer de forma sencilla y ajustada la información recogida. En el caso de datos cuantitativos hay ciertas medidas muy utilizadas como son la media, la mediana, la moda, la desviación típica o el rango de datos.

La última etapa, **el análisis de la información**. Su objetivo es determinar tendencias o patrones en el conjunto de los datos y extraer conclusiones.

5. CLASIFICACIÓN DE LAS VARIABLES

Los datos de los estudios se recogen en cuestionarios o formularios que permiten registrar la información de cada uno de los individuos de forma estandarizada y estructurada en forma de matriz de datos o tabla. Dicha tabla es una estructura que contiene los datos de cada individuo. Se estructura en filas y columnas de manera que la primera fila contiene toda la información del primer individuo, la segunda fila el del segundo individuo, etc. El número de filas coincide con el número de participantes del estudio de los que se han recogido datos. Las columnas representan las variables que se recogen, de manera que habrá el mismo número de columnas que número de variables.

Una variable es cada característica o aspecto estudiado en los sujetos de estudio que se puede observar y medir o caracterizar. Cada variable puede tomar distintos valores.

La estadística clasifica las variables en dos grandes grupos:

1. Variables categóricas o cualitativas: proporcionan datos no métricos.
2. Variables cuantitativas: proporcionan datos métricos.

A su vez se subdividen de la siguiente manera:

1. Variables categóricas o cualitativas.
   1.1. Variables categóricas nominales.
      1.1.1. Binarias.
      1.1.2. Con más de 2 categorías.
   1.2. Variables categóricas ordinales.
2. Variables cuantitativas.
   2.1. Variables cuantitativas discretas.
   2.2. Variables cuantitativas continuas.

Las **variables categóricas o cualitativas** no se pueden medir numéricamente. Variables cualitativas son la nacionalidad, el grupo sanguíneo, el sexo, el color de la piel, el color del pelo, la ocupación, etc.

Las *variables categóricas nominales* proceden de una escala de medida nominal. Son datos que, por su naturaleza, no admiten un orden entre los valores. Se pueden clasificar a su vez en binarias, con 2 categorías, como, por ejemplo, el sexo (masculino y femenino) y con más de 2 categorías, como, por ejemplo, la nacionalidad, el grupo sanguíneo o el título universitario (farmacia, enfermería, periodismo, ingeniería, economía, etc.).

Las *variables categóricas ordinales* tienen una escala de medida ordinal u ordenada y se puede establecer una jerarquía entre los valores. Por ejemplo: el nivel socioeconómico o el nivel de estudios de una persona.

Las **variables cuantitativas** vienen de una escala métrica, o sea, se pueden medir, tienen un valor numérico. Variables cuantitativas son la concentración de un agente químico, el nivel de presión acústica de un ruido, la edad de una persona, la presión arterial diastólica, etc.

Las *variables cuantitativas discretas* son aquellas que toman valores enteros y parten de recuentos. Por ejemplo, el número de veces que una trabajadora se ha infectado de un agente biológico. Los valores de esta variable serán, 1, 2, 3, pero nunca será 1,5. O el número de EPI que ha utilizado un trabajador en un año.

Las *variables cuantitativas continuas* son aquellas que proceden de una medición y pueden tomar cualquier valor dentro de un intervalo. Por ejemplo, la temperatura corporal (36,5ºC, 36,6ºC, 36,7ºC, etc.), el caudal de una campana de flujo laminar o la concentración ambiental de un agente químico en el trabajo.

Las variables cuantitativas permiten el cálculo de índices estadísticos como la media u otras medidas que veremos en los siguientes apartados. Se pueden hacer operaciones matemáticas con sus valores como sumas o restas. En cambio, las variables categóricas ordinales no permiten este tipo de operaciones, aunque estén codificadas con números. Por ejemplo, la piel se puede clasificar en función de su color, su capacidad para broncearse o la frecuencia de quemaduras solares en fototipos que van del 0 al 6. No se pueden hacer operaciones matemáticas con estos valores porque solo indican una escala ordinal, o sea, una posición relativa del individuo respecto al grupo.

6. ESTADÍSTICA DESCRIPTIVA

La estadística descriptiva es la rama de la estadística que se encarga de recolectar, analizar y caracterizar un conjunto de datos con el objetivo de describirlo y resumir sus características. Habitualmente se describen los datos de una muestra.

En la estadística descriptiva se utilizan parámetros o valores numéricos que ayudan a organizar y describir la información. Los parámetros pueden ser:

1. De centralización.
2. De dispersión.
3. De posición.

Los **parámetros de centralización** son aquellos valores en torno a los cuales se agrupan los datos. Los tres indicadores más habituales son la media, la mediana y la moda.

La *media* aritmética, o simplemente media, es el promedio de todos los datos. Su cálculo se realiza sumando todos los datos y dividiendo el resultado entre el número total de datos.

La *moda* es el dato de la muestra o del conjunto que más se repite. Puede ocurrir que haya dos números que se repitan el mismo número de veces y entonces se considera que la distribución de los datos es bimodal. También puede ocurrir que sean más de dos datos los que tengan la mayor frecuencia y eso muestra una distribución multimodal. Por último, puede que no haya moda porque todos los datos se repitan el mismo número de veces.

La *mediana* es una medida central que identificamos al ordenar los datos de menor a mayor. Cuando el número de datos es impar, la mediana es aquel dato que ocupa la posición central. Cuando el número de datos es par, la mediana es la media de los dos datos que ocupan las posiciones centrales. En definitiva, la mediana es el valor central del conjunto ordenado de observaciones. El 50% de los individuos de la distribución tienen valores inferiores a ella y el otro 50% tienen valores superiores.

Los **parámetros de dispersión** son aquellos que indican el mayor o menor grado de concentración de los datos alrededor de los parámetros de centralización. Al calcular la media de un conjunto de datos nos da una medida central, pero es importante conocer cómo se distribuyen el resto de los datos en torno a ella y estos parámetros lo describen.

Las medidas más habituales de dispersión son: la desviación respecto a la media, la desviación media, la varianza y la desviación estándar.

La *desviación respecto a la media* nos habla de cuanta diferencia o distancia hay de los datos del conjunto respecto a la media. A mayor desviación, más alejados los datos del conjunto de la media, a menor desviación, más próximos a la media.

La *desviación media* es la media de todas las desviaciones respecto a la media. Su cálculo se hace sumando todas las desviaciones y dividiéndolas entre el número de datos. Cuando la desviación media es muy grande los datos difieren mucho de la media, y si es pequeña apenas difiere de la media.

La *varianza* es una medida que representa la variabilidad de los datos respecto a su media y la *desviación típica o estándar* es la raíz cuadrada de la varianza.

Los **parámetros de posición** son aquellos que dividen a los datos en partes proporcionales de forma que cada parte tenga el mismo número de elementos. Para poder determinar los parámetros de posición es necesario ordenar los datos de menor a mayor. Los más habituales son: los cuartiles, los deciles y los percentiles.

Los *cuartiles* son tres valores que dividen al conjunto ordenado de menor a mayor en cuatro partes iguales. De manera que el primer cuartil es el que tiene el 25% de los datos detrás de él y el 75% de los datos por delante. El segundo cuartil coincide con la mediana y tiene el 50% de los datos por delante y el 50% de los datos por detrás. Y el tercer cuartil tiene por detrás el 75% de la distribución y por delante el 25% restante.

Los *deciles* son nueve valores que dividen el conjunto ordenado en diez partes iguales. De esa manera corresponderían al 10%, 20%, 30%, etc. de los valores. Y se llaman primer decil, segundo decil, tercer decil y así consecutivamente.

Po último, los *percentiles* dividen la muestra ordenada en cien partes iguales. En el caso de estar en el percentil 43, por ejemplo, el valor tendría el 43% de los datos por detrás de él y el 57% de los datos por delante de él.

7. ESTADÍSTICA INFERENCIAL

La estadística inferencial es la rama de la estadística que se encarga de determinar las propiedades de una población estadística a partir de la inducción o la inferencia. Extrae conclusiones sobre una población deducidas o basadas en la información que proporciona una o varias muestras de la población.

De la misma manera que la estadística descriptiva determina valores de la muestra, la estadística inferencial lo que trata es de extender esa información de la muestra a la población general.

Este área de conocimiento es muy útil cuando el equipo investigador solo tiene acceso a una o varias muestras de la población, pero no a la población completa. Por ejemplo, se ha hecho una

encuesta o una medición en una parte de un colectivo laboral, pero no en el colectivo completo. Dichas muestras deben ser representativas de la población de estudio para que las conclusiones de la inferencia sean correctas.

Hay dos grandes categorías en las que se clasifican los métodos de inferencia:

1. Métodos para la **estimación de parámetros poblacionales**. Volviendo al ejemplo anterior si el equipo investigador tiene acceso solo a una parte de la población de enfermería, por ejemplo, de España, pero no tiene acceso a todos los profesionales de enfermería necesitará hacer las mediciones sobre la parte conocida, sobre la muestra y después inferirlo a la población completa. Puede hacerlo como una estimación puntual o como una estimación por intervalos.

   Los *métodos puntuales* solo estiman un valor determinado. Por ejemplo, calculan el estadístico concreto de la muestra, como puede ser la media de edad de una muestra, y se toma como el valor más plausible o aproximado a la media de edad de la población general de enfermería.

   También existen los *métodos por intervalo* que determinan un margen de la muestra donde hay más probabilidad de que se encuentre el valor buscado de la población.

   La ventaja de utilizar la estimación por intervalo respecto a la estimación puntual es que habrá más posibilidades de dar la cifra correcta de la edad media de la población basándose en el estadístico calculado en la muestra. Como el intervalo marca un límite inferior de esa edad media y un límite superior, el valor de la edad media de la población completa es mucho más probable que se encuentre dentro de ese intervalo qué que coincida con el valor puntual de la muestra poblacional.
2. Métodos para **contraste de hipótesis**. Siguiendo con el ejemplo anterior en el que el equipo de investigación solo tiene acceso a una parte de la población, a una muestra de la población de enfermería, se mide una determinada propiedad de la población. La manera de hacerlo es a través de la muestra, puesto que no hay acceso a la población general de enfermería. Y tras hacer el cálculo en la muestra del parámetro concreto, por ejemplo, el grado de exposición a un determinado factor de riesgo, se formula la hipótesis de si el valor establecido en la muestra será igual o no al de la población general.

   Los métodos de contraste de hipótesis también se utilizan en las pruebas de independencia para corroborar la homogeneidad de varias muestras.

8. TÉCNICAS DE MUESTREO

Las organizaciones públicas y privadas necesitan, cada vez más, conocer información de los colectivos. Pueden ser sus necesidades, sus características, su comportamiento o sus opiniones. Analizar o estudiar a la población completa puede resultar muy costoso en términos económicos, temporales y humanos, por lo que es necesario buscar otros medios para acceder a la información sin necesidad de estudiar a cada uno de sus miembros. Por eso es necesario seleccionar una o varias muestras representativas que caractericen a la población general y que se puedan analizar para extraer resultados representativos de la población de referencia.

Las técnicas de muestreo son un conjunto de procedimientos orientados a estudiar y seleccionar un grupo de individuos que represente a la población. El objetivo es que la muestra represente

con alta probabilidad los parámetros de la población que se desean conocer, de manera que los resultados o estadísticos que se calculen en la muestra puedan extrapolarse al conjunto de la población. Esto implica que las conclusiones obtenidas de la muestra se considerarán válidas para toda la población y se generalizarán sus resultados.

Las técnicas de muestreo se pueden clasificar en dos grandes grupos:

1. El muestreo aleatorio o probabilístico.
2. El muestreo no aleatorio o no probabilístico.

El **muestreo aleatorio o probabilístico** es un proceso que permite obtener una muestra representativa de una población a través de un método de elección que se basa en la probabilidad de elección de los individuos de la muestra. En cambio, el muestreo no aleatorio no busca obtener una muestra representativa de la población. No se basa en un método de elección probabilístico, sino que la selección de los individuos de la muestra se determina por determinadas características y, por tanto, no todos los individuos tienen la misma probabilidad de ser seleccionados.

El muestreo aleatorio es ampliamente utilizado en el método científico porque presenta varias ventajas:

- Es el único que permite hacer los análisis confirmatorios de las hipótesis de partida y la inferencia estadística, o sea, extrapolación de los datos de la muestra a la población general confirmando los datos.
- Reduce los sesgos puesto que hay una probabilidad conocida de escoger a un determinado individuo de la población de referencia.
- Permite utilizar muestras de pequeño tamaño respecto a poblaciones grandes.

El proceso de elección de muestra tiene los siguientes pasos comunes:

1. La selección de la población de estudio. En este paso es necesario obtener el máximo de información de la población, como, por ejemplo, sus variables sociodemográficas como la edad, el sexo, la ocupación, el nivel formativo o el sector de actividad donde trabaja.
2. Elección de una técnica de muestreo aleatorio que dependerá de las características de la población.
3. Cálculo de la muestra mínima necesaria en función del tamaño poblacional.

Es importante señalar que cuando se selecciona la muestra se puede obtener representatividad de ciertas características de la población, pero no de otras, y esto es necesario conocerlo a la hora de hacer los análisis estadísticos posteriores.

Los tipos de muestreo aleatorio más comunes son:

Muestreo aleatorio simple. Se trata de asignar un número aleatorio a cada individuo de la población y después seleccionar una muestra de forma automatizada. Habitualmente se usa el método de la lotería, eligiendo un número al azar, o a través de un software de generación de números aleatorios.

Muestreo sistemático. Al igual que en el muestreo simple se enumeran todos los individuos de la población y se elige el primer elemento de la muestra de forma aleatoria. Se establece una

periodicidad regular de selección y se eligen los siguientes individuos de la muestra atendiendo a dicha periodicidad. Por ejemplo, tras asignar un número a cada individuo de la población se escoge uno en intervalos de 10: el décimo, el vigésimo, el trigésimo, etc.

Muestreo estratificado. Se parte de una población heterogénea que puede separarse en grupos homogéneos generalmente excluyentes, atendiendo a una característica, por ejemplo, la edad o el sexo. Una vez separados los grupos se realiza el muestreo aleatorio simple explicado anteriormente en cada uno de los grupos.

Muestreo por conglomerados o muestreo por racimos. Este método se suele utilizar cuando la población está muy dispersa, por ejemplo, geográficamente. Se trata de identificar bloques, conglomerados o *clusters* de población elegidos al azar de entre toda la población. La división podría ser en pueblos, ciudades, centros de salud, escuelas u otra característica similar. Una característica básica en este método es la existencia de heterogeneidad de los grupos. A su vez los elementos dentro de cada grupo deben ser tan heterogéneos como en la población de referencia.

El **muestreo no aleatorio o no probabilístico** es una técnica de muestreo en la cual el equipo investigador selecciona muestras basadas en un juicio subjetivo. La selección no se hace al azar ni obedece a la probabilidad. Eso quiere decir que no todos los elementos de la población tienen la probabilidad de ser elegidos en la muestra.

El objetivo y la utilidad de este tipo de muestreo es diferente a la del aleatorio. Se utiliza mucho en estudios exploratorios como las encuestas piloto previas a la encuesta definitiva donde el objetivo no es obtener datos representativos de la población sino comprobar la comprensión de las preguntas, identificar errores en el registro de los datos, cotejar los problemas de distribución de la encuesta, comprobar el tiempo de cumplimentación de la misma, etc. En este tipo de muestreo, el equipo toma decisiones sobre la muestra basándose en los conocimientos adquiridos durante la investigación. Se utiliza frecuentemente, aunque no únicamente, en investigación cualitativa.

Existen distintos tipos de muestreos no probabilísticos:

El *muestreo por conveniencia*. Se trata de una técnica de muestreo no probabilística donde la elección de la muestra depende de la disponibilidad de los sujetos para el equipo investigador. Son personas fáciles de reclutar, que están accesibles o disponibles, que no suponen mucho costo para el proyecto, y, desde luego, no son representativos de la población. En muchas ocasiones son conocidos o voluntarios.

El *muestreo consecutivo*. Es una técnica no probabilística muy similar al muestreo por conveniencia. En este caso el equipo investigador selecciona a un sujeto o grupo pequeño de sujetos analizando un solo tema durante un tiempo acotado. Una vez saturado y analizado ese tema pasa a otro sujeto o grupo de sujetos tratando un segundo tema diferente al primero. Es una técnica que ayuda al equipo investigador a trabajar muchos temas afinando la investigación a medida que va avanzando en el análisis de resultados de los primeros temas. En el planteamiento de la investigación puede cronológicamente plantear los temas en un orden lógico de entendimiento.

El *muestreo por cuotas*. En esta técnica se divide la población en partes, grupos o estratos en función de alguna característica o cualidad. La muestra no se basa en probabilidades sino en el conocimiento del equipo investigador de la población. Es similar al método probabilístico de

muestreo estratificado con la diferencia de que en el muestreo por cuotas los elementos de la muestra no se seleccionan aleatoriamente en cada estrato como sí ocurre en el anterior.

El *muestreo intencional o por juicio*. En este caso el único parámetro de selección del elemento de la muestra es el conocimiento y credibilidad del equipo investigador. El equipo elige solo a aquellas personas que consideran adecuadas para participar en el estudio de investigación.

El *muestreo de bola de nieve*. La técnica de bola de nieve se utiliza especialmente cuando es difícil para el equipo investigador localizar participantes para la muestra. Esto ocurre especialmente cuando el tamaño poblacional es pequeño y no está disponible con facilidad. Por ejemplo, un colectivo de profesionales muy específico o que se dedican a una actividad muy concreta o que tienen una enfermedad profesional muy rara. Una vez localizado el primer sujeto se le pide ayuda para encontrar sujetos similares con el objetivo de encontrar una muestra de suficiente tamaño.

Terminada la etapa de muestreo es necesaria la obtención de los datos. Las metodologías de obtención de datos son numerosas y variadas y dependen del diseño del estudio. Pueden ser las tomas de muestra sobre individuos, por ejemplo, análisis de muestras biológicas; o sobre ambientes de trabajo, como la determinación de la concentración de un contaminante o agente, etc. Sobre los datos obtenidos se llevará a cabo el análisis estadístico.

9. ESTADÍSTICOS PARAMÉTRICOS Y NO PARAMÉTRICOS

A partir de los datos observados y recogidos por el equipo investigador de la muestra se pueden calcular estadísticos para la muestra y parámetros para la población de referencia. Cuando se habla de los estadísticos paramétricos se parte del supuesto de que los datos vienen de una distribución normal. En cambio, los estadísticos no paramétricos no se basan en dicha suposición.

Las pruebas paramétricas analizan los elementos de la muestra basándose en las leyes de la distribución normal. Se aplican a variables numéricas y es necesario que la muestra elegida sea suficientemente grande para permitir un cálculo lo más exacto posible.

Existen diversas pruebas paramétricas como el cálculo del coeficiente de correlación de Pearson, el análisis de varianza factorial (ANOVA) o la regresión lineal. Entre las pruebas no paramétricas más conocidas: la prueba Chi (o Ji) cuadrado y los coeficientes de correlación e independencia para tablas cruzadas.

10. LA INVESTIGACIÓN CUALITATIVA Y SUS MÉTODOS

La investigación cualitativa es un conjunto de métodos y técnicas que intentan aproximarse al conocimiento de la realidad social a través de la observación y estudio de los fenómenos, del estudio de discursos, las acciones o de los símbolos sociales. Permite identificar las cualidades o atributos de un proceso social o un fenómeno, sus dimensiones, características, haciendo aflorar el discurso social desde el punto de vista de las personas que lo viven. El objetivo de la investigación cualitativa es comprender los significados de los fenómenos.

La investigación cualitativa detecta prejuicios y estereotipos y posibilita la extracción de resultados provenientes de estudios cuantitativos, complementando de este modo la comprensión de la realidad social.

La investigación cualitativa debe ser un proceso estructurado con una serie de elementos o componentes definidos de antemano. Es básico que el equipo investigador diseñe la investigación antes de comenzarla, y el primer paso es comenzar por el propósito de la propia investigación enmarcándolo en un tema concreto y explicando las líneas maestras que seguirá.

A continuación, se muestran los pasos del diseño de una investigación cualitativa. Cada investigación debe adaptarse en función sus propósitos y diseño:

1. **Definición del propósito de la investigación**. En esta fase se le pone título al proyecto, se explica el problema subyacente a la necesidad del estudio y el fin último que persigue el estudio.
2. **Exposición de los objetivos** generales y específicos del propósito de la investigación.
3. **Determinación de las preguntas** generales y específicas de la investigación.

4.1. **Abordaje del marco conceptual deductivo**. El marco conceptual o teorético es la recopilación, sistematización y exposición de los conceptos fundamentales para el desarrollo de una investigación. Al ser deductivo se parte de un marco conceptual que articula y hace operativos conceptos teóricos de otras teorías y paradigmas, o sea, constructos teóricos de otras autorías ya desarrolladas y probadas en investigaciones anteriores. En esta fase se definirán conceptos específicos, categorías dentro de dichos conceptos, indicadores o códigos para cada categoría y sus descripciones.

4.2 **Establecimiento del paradigma inductivo**. Tras la fase anterior donde el equipo investigador ha podido situarse y observar la realidad se pueden elaborar generalizaciones empíricas y construir teorías permitiendo encontrar estructuras más profundas para universalizarlas. En esta fase se determinan cuáles son las condiciones en las que se produce el problema, después se observan las estrategias de acción que realizan los actores de la situación y las consecuencias de dichas acciones.

Los pasos 4.1 y 4.2 pueden ser independientes si se opta por la deducción o la inducción o se pueden combinar. El uso individual o conjunto dependerá del problema a tratar y de las decisiones del equipo investigador. La aproximación deductiva es el diseño cualitativo más habitual en el campo de la salud laboral y la prevención de riesgos laborales.

5. **Diseño de la recogida y registro de los datos**. Esta fase del proceso implica trazar un plan de acción de recogida de datos. El equipo debe resolver cuál es la técnica o técnicas más adecuadas en función del objetivo y el resto de las características del estudio. Se puede optar por realizar entrevistas individuales, grupos focales, encuestas, etc. En esta etapa se contacta con los participantes, se realizan guiones de preguntas, se moderan las intervenciones, se graban las conversaciones, se transcriben las respuestas, etc.

6. **Análisis de los datos**. Una vez que se dispone de toda la información transcrita es necesario gestionar el amplio volumen de datos con la finalidad de organizarla y prepararla para la extracción de conclusiones. Las acciones en esta etapa implican realizar estrategias para analizar los significados como la organización, categorización, codificación y triangulación de la información.

7. **Establecimiento de las conclusiones** con el fin de probar o confirmar los resultados obtenidos.

8. **Determinación de las evidencias** de la investigación. Por último, el proceso determinará si existen evidencias que responden a la pregunta de investigación establecida al inicio del proceso y si, fruto de la investigación, han surgido nuevos conceptos.

En el apartado de diseño de toma de datos el equipo investigador debe determinar las técnicas de investigación cualitativa que va a utilizar. A continuación, se nombran algunos de los más comunes:

La **entrevista** es una de las técnicas más utilizadas en la investigación cualitativa. Hay diversas clasificaciones de las entrevistas en función del grado de libertad del equipo investigador, del tiempo disponible, del número de personas involucradas, etc. Las siguientes son las más comunes:

- La **entrevista individual** se realiza a una solo persona. Habitualmente se trata de una o varias personas entrevistadoras y una única persona entrevistada.
- La **entrevista grupal** se realiza a un grupo de participantes. El objetivo de este tipo de entrevista es conocer las opiniones de las personas que conforman el grupo. Suele seguir un guion prediseñado, por lo que se considera semiestructurada.
- La **entrevista en profundidad** tiene como finalidad explorar y entender la experiencia vital de otras personas y comprender el significado que ellas mismas le otorgan. Se puede realizar en una sola sesión o programar varias sesiones consecutivas que ayudan a aumentar el grado de profundidad de la conversación y permiten, a su vez, al investigador analizar los datos de las primeras sesiones para dirigir la exploración o las preguntas hacia el objetivo del estudio.
- El objetivo de la **entrevista estructurada** es conocer una determinada información a través de un mismo guion de preguntas, siguiendo el mismo orden y formulación de las mismas y aplicando escalas de valoración a todas las personas entrevistadas. Por lo tanto, en este caso es importante la preparación de la lista de preguntas. Se usa cuando hay muchas personas a las que entrevistar, se desea evitar la concurrencia de opiniones sesgadas, se requiere cierta rapidez y se necesita comparar de forma fácil y ágil los resultados de las diferentes personas entrevistadas. Es un proceso, por ejemplo, ampliamente utilizado en la selección de personal para un puesto de trabajo.
- La **entrevista no estructurada** no sigue ningún patrón estandarizado de preguntas siendo exclusiva para cada persona entrevistada. Por lo tanto, no tiene un guion previo de preguntas como sí es necesario para los otros tipos de entrevistas, aunque sí un esquema de temas. Se trata de un proceso muy dinámico y fluido que depende, en gran medida, de la habilidad de la persona entrevistadora. La entrevista se va construyendo a medida que avanza y en función de las respuestas que se obtienen del entrevistado.
- La **entrevista semiestructurada** permite al equipo investigador una cierta flexibilidad sobre el guion. Puede ser por poder cambiar el orden de las preguntas o porque pueda introducir preguntas nuevas que no estén incluidas en el guion. Habitualmente el equipo prepara una serie de temas o preguntas abiertas y a lo largo del proceso decide si las hace todas o no. En estos procesos se permite la realización de preguntas adicionales.

El **grupo focal** es una técnica de recolección de datos a través de una entrevista simultánea semiestructurada de varias personas. El grupo es una representación de un conjunto homogéneo de individuos que van a generar un discurso sobre un tema concreto a través del debate y de la interacción. El objetivo es que surjan actitudes, sentimientos, creencias, experiencias y reacciones de los participantes. Además, se trata de obtener una multiplicidad de miradas y procesos emocionales sobre un tema concreto y en el contexto del grupo. Los datos

que se obtienen se basan en la interacción de los participantes. La dinámica que se establece entre los distintos miembros del grupo permite conocer más profundamente las experiencias y el conocimiento de la materia que se pretende investigar.

El **análisis documental** es uno de los procedimientos menos conocidos en investigación cualitativa pero empleado desde el inicio de las ciencias sociales. El material documental de estudio proviene de fuentes documentales públicas o privadas que pueden ser históricas, de archivo, literarias, imágenes, videos, audios, dibujos, en papel, informáticas, etc. El documento es cualquier manifestación humana que contenga un significado para la investigación y que no haya sido creada *ad hoc* para la investigación.

El ***shadowing*** es una técnica en la que hay, al menos, dos personas involucradas: la persona que observa y la persona observada. El observador sigue a la persona observada como si fuera su sombra. Lo hace durante un determinado periodo de tiempo y suele ser útil para entender la forma de uso de un producto o un servicio. También es muy útil en la observación de comportamientos durante procedimientos de trabajo. Es muy importante que la persona que observa no interfiera en el comportamiento de la persona observada. Es una técnica que puede durar media hora o varios días en función de la información que se quiera obtener y las características del proyecto. Es una técnica ampliamente utilizada en ergonomía.

El **análisis de las comunidades *online* o redes sociales**. Las comunidades en línea están formadas por participantes que comparten un interés común. En ellas se puede obtener mucha y variada información sobre productos, servicios, temas concretos. El objetivo de analizar este tipo de datos es reunir información detallada y cualitativa en grupos de personas que tienen mucho conocimiento sobre un tema durante un tiempo concreto. Estas comunidades proporcionan la base para nuevas ideas, innovaciones y formas de abordar un problema. Las conversaciones e intercambio de opiniones se suelen producir de forma natural y la persona que modera el foro solo suele tener que introducir un nuevo tema para que el resto de participantes expongan su conocimiento u opinión sobre el tema abordado.

Hay otros ejemplos de investigación cualitativa como son el estudio de diario, la investigación etnográfica, técnicas de consenso, participativas, etc.