METODOLOGÍA DE LA INVESTIGACIÓN

Planteamiento del problema: la pregunta de investigación

Gonzalo de la Fuente Echevarría

TODO BUEN ESTUDIO de investigación debería partir de una premisa: necesidad no satisfecha. Esa debería ser el resumen y el germen que termine desembocando en un estudio de investigación. Sin embargo, esto requiere muchos factores que deben asociarse como "nutrientes" de esa necesidad, ya que los pilares que la sustentan, a nivel práctico, son una actitud crítica, escéptica; una actitud escéptica ante la práctica habitual asignada como dogma de fe sin comprobar la evidencia en la que se sustenta. Es autoconsciencia de que los conocimientos actuales que practicamos no lo serán el día de mañana, y en mayor o menor medida, puede depender de nosotros. Exige innegablemente una capacidad crítica de observación (escuchar y observar a nuestros pacientes, obvio, pero siempre vigente) que despierte la curiosidad ante hallazgos que no cuadren con los conocimientos aprendidos. Un paciente al que dar una respuesta a su problema médico y una ausencia de evidencia científica que permita pronunciarnos en su resolución. Esto, obviamente, exige una capacidad científica para buscar y discriminar en la literatura existente, como veremos más adelante, que cribe nuestras necesidades en "satisfechas o no". Otras veces el proceso puede iniciarse por un efecto deseado o no deseado y una necesidad de comprobar la implicación

de un factor o factores en su desarrollo. La exposición a un factor "x" y la necesidad de comprobar cómo puede afectar a unos determinados pacientes. Muchas son las causas que pueden originarlo si mantenemos nuestra inquietud intacta[1].

1-De la pregunta clínica a la pregunta de investigación

Como antes decíamos, todas estas posibilidades que se nos presentan a diario, deberían partir de esa insatisfacción para contestar nuestra duda a través de lo ya publicado. Esto implica diversos elementos que debemos conocer, ya que deberán contestarse de una manera secuencial como si de un algoritmo se tratara. Nuestro paciente es la raíz del algoritmo, y su problema debería germinar en una pregunta clínica.[2] Lo lógico es que ese paso desembocara en una búsqueda bibliográfica, con el objeto de esclarecerla. Esto implica un conocimiento exhaustivo de la metodología para hacer una pregunta clínica (como veremos más adelante, según la metodología PICO), para poder formular correctamente nuestra necesidad y plasmarla de una manera efectiva en una respuesta con la bibliografía existente. Se evalúa de forma crítica la evidencia y, si queda resuelta, se integra con la experiencia individual y se finaliza el proceso. Esto debería aplacar la mayoría de las iniciativas de investigar, ya que el hecho de encontrar una evidencia de mayor grado de la que podamos hacer nosotros en consulta abortaría la necesidad de gastar esfuerzo y recursos y no contribuir a la toxoinformación. En este sentido, conviene recordar la pirámide de la evidencia, que estructura los recursos de información de acuerdo con su utilidad y propiedades en la toma de decisiones en la atención sanitaria. Como ya sabemos, la pirámide contiene cinco escalones (también llamada "de 5S").

Si por el contrario la duda planteada no se resuelve, debería germinar en nosotros la necesidad de plantearnos evolucionar la pregunta, iniciando las fases de un proceso de investigación. Tal como procedimos para formular la pregunta clínica para contestar un "vacío propio" de información, la pregunta de investigación será la base para llenar un "vacío común" de información. Son la parte más importante de un estudio, los

cimientos, porque si la pregunta no está bien planteada, las conclusiones pueden ser erróneas, por mucho que la metodología sea correcta. Nacen de la misma raíz estructural, sólo que debemos ajustarla según una serie de variables. Como refleja JW Tuckey, "con diferencia, es mucho mejor una respuesta aproximada a una pregunta correcta, que una respuesta exacta a una pregunta errónea". El proceso de creación del proyecto de investigación habrá dado comienzo (Figura 1).

Figura 1 Características de la pregunta de investigación.

La formulación clara de una pregunta de investigación con nuestra necesidad de conocimientos nos ayuda además a identificar el tipo de diseño del estudio que tendrá mayor probabilidad de responderla, facilita el análisis de los resultados y la valoración de los mismos. Por lo tanto, debemos asegurarnos de que la pregunta sea pertinente

y lo más concreta posible. Plantear una pregunta de investigación requiere lograr un equilibrio entre una pregunta de gran amplitud que puede resultar en una carencia de dirección o el análisis superficial de un área extensa, y una pregunta específica que conduzca a un análisis exhaustivo de un problema bien delimitado. Es la base del proceso de adquisición de conocimientos con evidencia científica y lo será de los estudios que decidamos hacer en caso de no encontrar dichas respuestas. Como explica Juan Bautista Cabello: "la formulación de preguntas es una habilidad clínica fundamental, al ser un instrumento de conexión entre la práctica clínica y el conocimiento en los dos sentidos: aplicar conocimiento a la práctica y generar desde la práctica preguntas para la investigación".

Como decíamos, la pregunta debería ser pertinente. Esto nos debe generar ciertos campos que debemos comprobar o ajustar para que se adecue nuestra pregunta, incluidas en los criterios FINER (*factible, interesante, novedosa, ética, relevante*). Debe ser ***Factible***, es decir, tiene que poder ser respondido. Si no tenemos los recursos o el proyecto escapa de nuestras posibilidades, no será pertinente. Además, debe ser ***Interesante***, es decir, debe aportar una diferencia en nuestro proceder médico o conocimiento, que pueda modificar o reafirmar nuestra práctica habitual. La pregunta de investigación debe ser ***Novedosa***, ya que, como decíamos anteriormente, debe estimular la búsqueda de respuesta a nuevas preguntas, a conocimientos aun no adquiridos. Decíamos que debe ser además Ética, ya que, si queremos encontrarla en la bibliografía o va a generar un estudio, debe poder ajustarse a la práctica médica, sobre todo en lo referente a estudios experimentales. Por último, pero obviamente no por ello menos importante, debe ser ***Relevante***, ya que debe poder ser aplicable a nuestra práctica o a la de los demás, ya que de nada sirve encontrar una respuesta que sea imposible llevar a cabo. La relevancia se puede predecir si imaginamos los distintos resultados de la investigación y consideramos cómo pueden

afectar al conocimiento científico, a la práctica clínica y a la dirección de investigaciones futuras.

2-Formulación de la pregunta de investigación.

En algunos casos, el paso de la idea al planteamiento del problema puede ser inmediato, casi automático, o bien llevar una considerable cantidad de tiempo; dependerá de la experiencia del investigador, la complejidad de la idea, la existencia de estudios anteriores, el enfoque elegido, el empeño del investigador y las habilidades particulares.

Además de la construcción de la pregunta de investigación, podemos realizar un esbozo del plan de estudio de una o dos páginas, lo que puede ayudar al investigador a clarificar sus propias ideas sobre el plan, concretar los objetivos del estudio y a descubrir problemas que requerirán nuestra atención. Aquí plasmaremos con antelación las limitaciones de nuestro estudio, y valorar si requeriremos ayuda de investigadores con más experiencia.

Una vez detallas, en el punto anterior, las características que debería integrar nuestra pregunta de investigación, podemos pasar a explicar la manera por la cual la que podemos llevarlo a cabo. La aproximación es el uso de una «sintaxis estructurada» propuesta por Richardson y que hemos llamado estructura PICO (o su variante ampliada PICOT), por la nemotecnia usada en inglés (**Patient, Intervention, Comparison, Outcome +/- Type or Time**) (Figura 2).

Si desglosamos cada punto de la pregunta de Investigación, entenderemos como debemos dar los pasos a seguir para generar la pregunta, ya que, como vimos antes, es muy importante el "cómo" plasmar una duda.

1. **Paciente:** El "objeto enfocado" al que queremos medir el efecto de someterle a una intervención. por ello debemos definir, lo más detalladamente posible, si la intervención y nuestra capacidad técnica lo permite.

2. **Intervención:** Es el factor que nos generó la duda en nuestro paciente. Aún más importante que este bien definido, ya que la comparación y resultados van a depender de que la intervención sea clara. Debemos ser capaces de diferenciar y medir el cumplimiento de ésta para después recoger los resultados. Si no podemos controlar cómo se realiza la intervención (en los estudios experimentales), o cómo se ha registrado la exposición producida (en los estudios observacionales), los resultados se verán claramente comprometidos.
3. **Comparación:** No siempre procede o no siempre es posible tenerlo. Cuando el resultado sea analítico debemos definir el factor de comparación, tanto si son factores de exposición como si son intervenciones. La pregunta cambia claramente su enfoque dependiendo de la comparación seleccionada.
4. **Resultado/Outcome:** Siempre tendemos a pensar en este como el elemento clave, ya que representa aquello que queremos comprobar. Es un elemento rico en matices, conviene "pulirlo", definirlo al máximo, para dar validez a nuestros resultados. Los matices van a condicionar que el resultado sea, a grandes rasgos, positivo o negativo pareciendo la misma pregunta.

Algunos autores incluyen aquí un 5º factor, que designaríamos como "T": Tipo de estudio y/o tiempo necesario para medir el resultado. Representaría el tiempo necesario para el estudio y qué diseño es el más adecuado para contestar la pregunta. Este punto podrá definirse mejor al determinar la metodología a seguir, lo que tiene que ver con los objetivos propuestos (como veremos más adelante).

Existe alguna alternativa, como el *método sistemático de Bordage y Dawson*, pero no lo incluiremos para no resultar más complejo. Recomendamos practicar esta sistemática y aquellos investigado-

Figura 2. Estructura PICO o PICOT

res que resulten frustrados por ésta, simplemente conozcan que existen otras alternativas.

Conviene destacar, por si alguien está familiarizado, que algunos autores consideran realizar en este momento la búsqueda de información (y no antes), con el fin de no condicionar la creación de su pregunta de investigación en función de la bibliografía previa. De esta forma, se crea la pregunta de investigación y después se recurre a las fuentes (primarias, secundarias...), de tal modo que se realice una "depuración de la pregunta". Así se puede ver si es un tema estudiado, ya investigado y requiere dilucidar algún matiz o unificar criterios en

un único estudio, o aplicar el estudio a una población diferente, etc. Se encuentran ser múltiples motivos en los manuales de investigación, aunque recomendamos realizar la búsqueda previa, para adquirir todos los conocimientos o resolverlos con la literatura ya publicada, y solo investigar aquellos que, como decíamos, no puedan ser resueltos tras ésta.

Aunque en muchos estudios pueden incluir más de una pregunta de investigación, es importante intentar establecer una única pregunta cuando se diseña el trabajo. Esta pregunta nos ayudará a redactar el objetivo de nuestro estudio (qué se pretende conocer, en qué población y en qué contexto), qué posible relación pensamos que puede haber entre las variables que se estudian (hipótesis) y, al menos *a priori*, qué tipo de estudio debemos realizar para conocer lo que se busca responder.

Por todo ello, si bien podría admitirse que no todos los pediatras que trabajan en Atención Primaria aporten producción científica, sí que se debe asumir el compromiso de transferir los resultados de la investigación existente a la práctica y basarla en la mejor evidencia científica disponible. Por lo tanto, es preciso ser consumidores de investigación, consumidores con formación, inteligentes y críticos, con capacidad de discernir lo verdaderamente valioso y riguroso de aquello que no lo es o que tiene intereses no legítimos.

3- Clasificación de las preguntas de investigación

Vamos a describir someramente los posibles tipos de preguntas clínicas para entender como generarlas. Podemos clasificar las preguntas clínicas en diferentes categorías. Creemos que tener claras las categorías puede resultarnos útil para entender qué tipo de pregunta de investigación puede dar forma a esa necesidad de conocimiento. De esta forma, distinguimos tres categorías, según:

- La *intencionalidad* de la pregunta. En función de lo que el investigador pretende con la pregunta que está formulando se distinguen:
  - *Preguntas descriptivas*: describen un fenómeno de la naturaleza en un punto específico del tiempo y el espacio. Se centran en el primer paso del método científico, es decir, la observación. Por definición, no hay comparaciones ni se plantean hipótesis, aunque pueden servir de base para estudios futuros. Incluyen generalmente un adjetivo interrogativo (cuál, cuánto, quién..), la medición (prevalencia, incidencia), una condición (asma, diabetes…), la población, el lugar y el momento.
  - *Preguntas inferenciales o analíticas*: comparan intervenciones, técnicas o exposiciones para determinar su asociación con un desenlace. Las que más se ajustan a la metodología PICO. Ante estas preguntas se formulan hipótesis.
- La *finalidad* de la pregunta. Según el resultado específico esperado por el investigador pueden ser:
  - *Preguntas cuantitativas*: tratan la variabilidad de un aspecto clínico o epidemiológico.
  - *Preguntas cualitativas*: tiene como objetivo encontrar significados, interpretaciones o explicaciones de un fenómeno que no es pertinente o posible cuantificar. La finalidad, muchas veces, es generar nuevas hipótesis o modelos teóricos, que permitan desarrollar futuros estudios o la comprensión de ciertos fenómenos. No se suelen ajustar a la estructura PICO.
- El contexto clínico en el que la pregunta se encuentra inmersa. En la práctica clínica se diferencian cuatro actividades básicas:

- *Etiología o causalidad*: Estas preguntas tienen una connotación negativa. Aparecen cuando el investigador piensa en un factor que aumentará la probabilidad de sufrir una enfermedad o condición.
- *Diagnóstico*: son difíciles de formular. Con estructura de tipo descriptiva. El objetivo de estas preguntas es determinar la capacidad de una prueba para discernir correctamente si un paciente sufre o no una enfermedad según el resultado de un gold estandar.
- *Intervención*: son típicamente analíticas. Se evalúa generalmente la prevención o el tratamiento (farmacológico, terapia, estrategia diagnóstica...).
- *Pronóstico*: predicción de las consecuencias de la condición estudiada en el tiempo. Servirá para identificar a qué grupo de personas con una condición le irá mejor o peor en el futuro.

Formulación de la hipótesis y objetivos.

Gonzalo de la Fuente Echevarría

EL NEXO ENTRE la pregunta y el estudio de investigación lo constituyen las hipótesis de trabajo. Una pregunta correctamente estructurada conducirá a formular la hipótesis de investigación. La hipótesis es una oración declarativa que anticipa los resultados de un estudio de investigación basado en el conocimiento científico existente y en los supuestos declarados. No puede ser una mera conjetura desordenada o caótica, irregular, sino reflejar el conocimiento, imaginación y experiencia del investigador. Por lo tanto, sería también una "predicción controlada" que responde a la pregunta de investigación. 10

Una hipótesis bien formulada cuenta con una estructura compuesta por una unidad de observación (sujetos u objetos) y variables (atributos susceptibles de medición); además, se puede indicar cómo se espera que se relacionen estos dos elementos (direccionalidad de la hipótesis). Cabe destacar que esa direccionalidad de una hipótesis traduce las expectativas, lo cual, en ocasiones, puede ir en detrimento de su imparcialidad. No obstante, todo investigador tiene cierta idea o intuición sobre la posible respuesta a su problema, aunque no la formule explícitamente, de ahí que lo denominemos como "predicción".

Esa predicción se puede realizarse de una manera inductiva o deductiva. La primera, partiendo de una observación de un problema concreto que conduce a la formulación de una hipótesis general. El método deductivo, por el contrario, supone la extracción de resultados en base a una premisa que se considera como verdadera, es decir, parte de una ley universal, para determinar si se aplica a un caso particular.

La hipótesis no debería cambiarse una vez se obtienen los resultados. Si estos son discordantes, se pueden reflejar como tal en los resultados y ser la base de futuros estudios.

Hay dos tipos de hipótesis, la nula (no hay diferencia entre los dos grupos que se comparan) y la alternativa (hay diferencia y además en un determinado sentido, positivo o negativo). Una vez encontrada la asociación habrá que darle un sentido de causalidad (si se puede, siguiendo los criterios establecidos por Sir Austin Bradford Hill (Figura 1). De acuerdo con la hipótesis, que suele plantearse en su forma alternativa, se deciden los objetivos, es decir, la hipótesis se traslada a variables operativas que se pueden y deben medir. Todos los proyectos de investigación deben incluir la formulación directa y concreta de las intenciones y objetivos de la investigación. Los objetivos deben ser escasos, estar descritos de manera concisa y ser realizables (factibles). Así pues, el problema-pregunta precede a la hipótesis-respuesta que, a su vez, deriva del o de los objetivos de la investigación. Se debe diferenciar el propósito general del estudio, con las expectativas de la investigación y la relación entre las variables que analiza, de los objetivos específicos, que se desprenden del general, aunque englobados dentro de la línea de pensamiento de éste.

A la hora de elegir la medida principal deberíamos tener en cuenta su importancia clínica y la posibilidad de ser capaces o no de realizar su medición (por nuestros medios o por lo frecuencia del evento). Es decir, si el resultado que queremos medir requiere unos medios materiales de los que no disponemos, es tan subjetivo que no resulta reproducible o simplemente se produce en una frecuencia demasiado

Figura 3. Criterios de Causalidad de Austin Bradford Hill

baja para apreciarla y recoger suficiente número de eventos, comprometerá los resultados.

Es factible que una pregunta de investigación tenga varias medidas de resultado posibles. De manera inicial debe elegirse la medida principal de efecto, ya que, en función de la medida elegida, plantearemos el tamaño muestral necesario y los principales análisis. La existencia de medidas de efecto secundarias permite complementan el análisis y ayuda a evaluar la consistencia interna de los resultados.

Una vez hecho esto, el proceso nos dará pie a plantear qué tipo de estudio se adapta mejor para llegar a responder su pregunta. Es un proceso encadenado. Es decir, la pregunta de investigación, la hipótesis y el objetivo provocan una reacción en cadena que debería conducir "por si sola" al tipo de estudio. En esta decisión influirán la validez del diseño y los recursos disponibles (tiempo, información, población disponible, aspectos éticos, financiación económica, etc.). Si para

responder al objetivo hace falta un estudio no factible, deberá replantearse el objetivo o hasta "desechar" la pregunta de investigación. Como decíamos, teniendo la hipótesis clara, y los objetivos definidos, el investigador puede aproximarse a esa realidad "observando lo que sucede" (estudios observacionales) o manipulando y controlando la exposición para ver qué sucede (estudios experimentales). Cuanto más controle mayor fortaleza tendrán las conclusiones y viceversa, cuanto menos intervenga mayores posibilidades de sesgo tendrá el estudio.

Tipos de estudios de investigación

Gonzalo de la Fuente Echevarría

ENTENDEMOS POR DISEÑO de estudio el conjunto de procedimientos, métodos y técnicas mediante los cuales se actúa con los participantes del estudio, se recopilan los datos, se analizan los resultados y se interpretan para obtener las conclusiones. Como decíamos anteriormente, el tipo de diseño del estudio dependerá del tipo de problema que nos induce a realizar la investigación, y surge casi "de manera natural", cuando hacemos el trabajo previo.

La decisión de que estudio realizar es muy importante, porque influirá enormemente en el diseño, en los posibles sesgos que deberá intentar controlar (muchos en los estudios observacionales y pocos en los experimentales) y en la extrapolación de las conclusiones (los estudios experimentales, aunque tienen pocos sesgos y gran potencia estadística, a veces no son tan fácilmente generalizables porque la población general no siempre está bien representada en los sujetos estudiados tras unas rigurosas condiciones de inclusión). Es posible seleccionar un diseño de investigación completamente equivocado para responder una pregunta específica. Por ejemplo, es posible que desee responder a una de las preguntas de investigación descritas anteriormente: "¿Que páginas de divulgación medica consultan las familias?" Aunque muchos

consideran que un estudio controlado aleatorio es un diseño de investigación "gold estándar", dicho estudio simplemente no sería capaz de generar datos para responder la pregunta planteada.

El lenguaje de la pregunta de investigación puede ser útil para decidir qué diseño de investigación y métodos a utilizar. Por ejemplo, si la pregunta comienza con "cuántos" o "con qué frecuencia", probablemente se trate de una pregunta descriptiva para evaluar la prevalencia o incidencia de un fenómeno. Sería apropiado un diseño de investigación epidemiológica, tal vez utilizando una encuesta o entrevistas estructuradas para recopilar los datos. Si la pregunta comienza con "por qué" o "cómo", entonces es una pregunta descriptiva para obtener una comprensión profunda de un fenómeno. Un diseño de investigación cualitativa, utilizando entrevistas en profundidad o grupos focales, recopilaría los datos necesarios. Finalmente, el término "cuál es el impacto de" sugiere una pregunta causal, que requeriría la comparación de los datos recopilados con y sin la intervención (ensayo clínico).

Para escoger el estudio correcto que dé respuesta a nuestra pregunta de investigación, para contrastar a nuestra hipótesis y que cumpla nuestros objetivos, debemos conocer y entender las peculiaridades de cada tipo de estudio. Hay múltiples clasificaciones de los tipos de estudios disponibles. Hemos creído oportuno simplificarlo y clasificarlos según el objetivo general. Dejamos un gráfico con el concepto general de todos ellos (Figura 4). Veamos la clasificación de los que vamos a tratar con importancia en Atención Primaria, para después desglosarlos en profundidad:

Estudios transversales o de prevalencia

Son estudios observacionales y descriptivos, que carecen de direccionalidad (reflejan "una foto" de la situación estudiada). Tienen como objetivo genérico acumular datos para describir fenómenos aún poco conocidos. Como todas las variables se miden en el mismo momento, no es posible establecer relaciones temporales y, por tanto, de

causa-efecto, por lo que los estudios transversales son estudios de prevalencia y siempre de naturaleza descriptiva. Son estudios útiles para la planificación sanitaria, ya que informan de la distribución de enfermedades y de factores de riesgo, por lo que ayudan a formular hipótesis etiológicas que luego deberán ser comprobadas con otros tipos de estudios (analíticos). Son los diseños más comúnmente encontrados en las revistas científicas y en este tipo de estudios no existe un grupo de comparación. Por ello son, en general, más baratos en la planificación y con menos complejidad. Pueden ser útiles para plasmar una condición existente u objetivar una observación subjetiva que, a posteriori, nos de paso a planificar un estudio analítico.

Estudios analíticos

Pretenden poner en evidencia asociaciones causales e intentan averiguar el porqué de ciertas situaciones. En ellos se prueban las hipótesis planteadas y se requiere de datos que sostengan dichas respuestas. Encontramos dos tipos básicos:

Estudios observacionales

Estudios en los que los investigadores actúan como meros observadores ya que existen limitaciones éticas que impiden la manipulación del investigador. Por lo tanto, no controla la exposición al factor, sino que ya le viene dada. De ahí que la validez del estudio sea menor que la de los diseños experimentales. Dentro de éstos vamos a desglosar dos tipos básicos (Figura 1):10,12-17

Casos y controles:

De una determinada población se seleccionan dos grupos, unos con determinada enfermedad (casos) y otros sin ella (controles). Se les pregunta retrospectivamente sobre una determinada exposición para valorar si la exposición al factor es más frecuente en un grupo que en otro. Los casos deben estar perfectamente definidos y con unos criterios muy estrictos. Los controles, que se van a utilizar para estimar la prevalencia de exposición al factor en la población de la que provienen

los casos, pueden seleccionarse tanto de un medio hospitalario como de la población general, pero su selección no debe verse influida por el grado de exposición al factor de estudio. Las medidas que podemos calcular en éstos son la proporción de expuestos, tanto en casos (respecto del total de casos) como en controles, así como las proporciones de no expuestos, de manera complementaria a la anterior. Se puede calcular en ellos la odds ratio que nos indica, entre 0 y 1, de manera inversamente proporcional, que la exposición actúa de cómo factor protector a la enfermedad. Y si es mayor de 1, de manera proporcional a su valor, que la exposición supone un riesgo de enfermedad.

Suelen ser menos costosos y duraderos que los de cohortes y permiten, además, el estudio de varios factores de exposición para un mismo efecto, además de ser ideales para estudiar enfermedades raras. Pero los principales problemas que presentan son la alta susceptibilidad a presentar sesgos (principalmente selección e información o análisis), no se pude estimar directamente la incidencia de enfermedad, ni la secuencia temporal entre exposición y efecto. No permiten la estimación de incidencia ni prevalencia de enfermedad, como en el anterior, y por lo tanto tampoco de riesgos relativos.

Estudios de cohortes:

Son de tipo observacional, analíticos (hay grupo de comparación), habitualmente anterógrados y de temporalidad concurrente o mixta. Se seleccionan dos o más cohortes o grupos de personas en base a su exposición al factor. El grupo o cohorte sometida a un factor de exposición es seguida a lo largo del tiempo para comparar la frecuencia de aparición del efecto respecto a otra cohorte no expuesta, que actúa como control. Los sujetos son seguidos hasta que desarrollan el efecto, hasta que se pierden durante el seguimiento o hasta que finaliza el estudio. En ellos se pueden calcular los riesgos en expuestos y en no expuestos, para los estudios de incidencia acumulada. Con ello calcularemos las medidas de asociación, como son el riesgo relativo (RR), la reducción absoluta de riesgo (RAR) y la reducción relativa del riesgo

(RRR). No abordaremos este tema en profundidad, ya que se tratará más detenidamente en el tema siguiente.

Su principal ventaja es que permiten registrar la incidencia (casos nuevos que aparece en un periodo de tempo en la población) del efecto y la evolución de la enfermedad, por lo que permiten establecer hipótesis de cara a posteriores estudios experimentales. Además, tienen menor posibilidad de sesgos en la medición de la exposición que otros estudios observacionales. Entre sus principales inconvenientes están su elevado coste y dificultad de ejecución. Además, son poco útiles para estudiar enfermedades raras o con largos períodos de latencia y son susceptibles al cambio de las circunstancias a las pérdidas de participantes durante el seguimiento, ya que no se pueden controlar tanto las condiciones como en los experimentales (aunque también resultan mucho más baratos que éstos, sobre todo los retrospectivos).

Estudios Experimentales

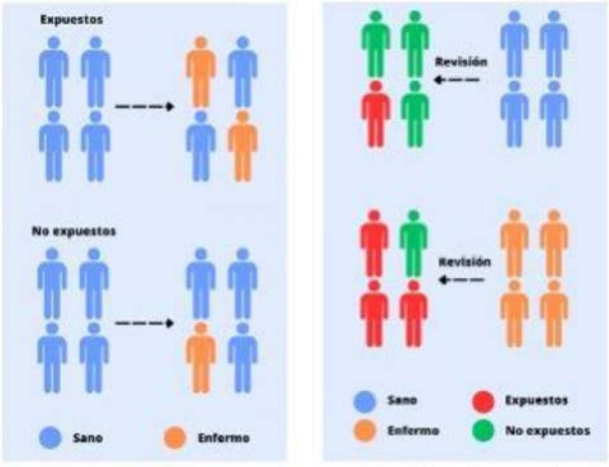
Ensayo Clínico Aleatorizado:

Son estudios analíticos, en los que se interviene sobre la exposición. Son por ello anterógrados, de temporalidad concurrente y de muestreo de una cohorte cerrada con control de la exposición. Este tipo de estudios permite la exposición controlada al factor para minimizar el riesgo de sesgos de otros estudios, además de obtener información más fiable sobre la relación causal entre exposición y efecto. Son los estudios que proporcionan una mayor seguridad sobre inferencia causal y los que tienen una mayor validez externa, además del menor riesgo de sesgos por la selección aleatoria de los grupos de intervención y control. Tras seleccionar a un grupo de sujetos que cumplen unos criterios de inclusión (características clínicas y sociodemográficas de los enfermos) se eliminan a los que pueden presentar alguna contraindicación para someterse a la intervención (criterios de exclusión). Se efectúa una medición de variables basales, para conocer las características de los sujetos (también sirve de ayuda para mejorar la elegibilidad); generalmente

después se aleatorizan, es decir, se reparten en dos o más grupos de tratamiento de una manera no predecible ni manipulada (generalmente mediante una secuencia o código de aleatorización). Esto permite, en teoría, homogeneizar los grupos de intervención y control distribuyendo de manera similar los factores de riesgo que puedan presentar los participantes del estudio (evitar el sesgo de selección). Esta aleatorización puede ser simple, por bloques, estratificada o mediante minimización. Este proceso de asignación de los participantes puede ser, siempre que sea ético, desconocido para los participantes (ciego simple), también para los médicos que administran el tratamiento (doble ciego) o incluso también para los analistas de los datos (triple ciego).

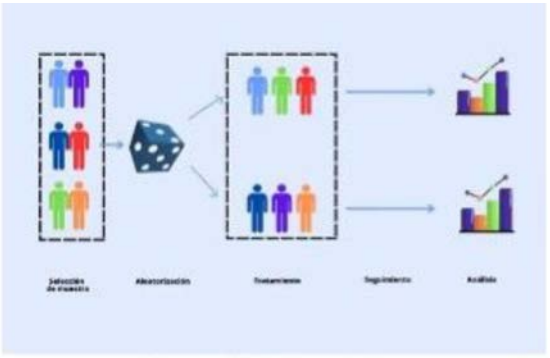
Se les sigue durante un periodo determinado (y lógico como para que se produzca un efecto) y posteriormente se analizan las diferencias de presentación del efecto en cada uno de los grupos. Todo esto facilita que sean los más reproducibles y comparables con los resultados de otros estudios. Entre sus inconvenientes, su coste en tiempo y recursos y los problemas éticos que pueden surgir al exponer a determinados factores de exposición. Además, la propia rigidez de la colección de participantes y de intervención pueden hacer en ocasiones hacer que sus resultados sean difíciles de generalizar, y no están exentos de sesgos.

Figura 1. Tipos de estudios



1.Estudios de Cohortes

2. Estudios de Casos y controles



3. Ensayo clínico aleatorizado

Bibliografía

BUÑUEL ÁLVAREZ JC, RUIZ-CANELA CÁCERES J. Como elaborar una pregunta clínica. Evid Pediatr. 2005; 1: 10.

OCHOA SANGRADOR C. Diseño y Análisis en Investigación. IMC, Madrid 2019.

ARGIMÓN PALLAS JM. JIMÉNEZ VILLA J. Métodos de investigación aplicados a la Atención primaria de salud. Barcelona: Harcourt; 2000.

CUMMINGS. Diseño de la investigación clínica. Un enfoque epidemiológico Barcelona: Ed. Doyma; 1993.

CABELLO LÓPEZ JB. Lectura crítica de la evidencia clínica. 2ª Edición. Ed. El Sevier 2021.

HAYNES B. Forming research questions. Clinical epidemiology: How to do clinical practice research. 3.a ed., pp. 496.

RUIZ. JG. La pregunta de investigación. Epidemiología clínica, investigación clínica aplicada. 1.a ed., pp. 29-50.

MARTIN C. La pregunta de investigación en la práctica clínica: guía para formularla. Revista Colombiana de psiquiatría. Vol 47. Num 3, pp 193-200.

TULLY MP. Research: articulating questions, generating hypotheses, and choosing study designs. Can J Hosp Pharm. 2014;67(1):31-4.

MARTINEZ SUÁREZ V. Investigar en Atención Primaria. Pediatr Integral 2012; XVI(8): 663.e1-663.e7

ICART MT. El uso de hipótesis en la investigación científica. Rev Atención Primaria. 1998. Vol 21. Num 3. pp 172-178.

MOLINA ARIAS M. OCHOA SANGRADOR C. Tipos de estudios epidemiológicos. Evid Pediatr. 2013;9:53.

MANTEROLA C, QUIROZ G. Metodología de los tipos y diseños de estudios más frecuentemente utilizados en investigación clínica. Rev Médica Clínica Las Condes. 2019. Vol 30;1:36-49.

MOLINA ARIAS M, OCHOA SANGRADOR C. Estudios Observacionales (II). Estudios de cohortes. Evid Pediatr 2014;10:14.

MOLINA ARIAS M, OCHOA SANGRADOR C. Estudios Observacionales (III). Estudios de casos y controles. Evid Pediatr 2014;10:33.

MOLINA ARIAS M, OCHOA SANGRADOR C. Ensayo clínico (II). Resultados. Variables. Medidas de impacto. Evid Pediatr 2015;11:33

Aspectos metodológicos básicos del estudio

Venancio Martínez Suárez
Daniel Mata Zubillaga

Selección de la muestra

DESPUÉS DE DEFINIR el tipo de estudio deberemos contestar a un interrogante: ¿Quiénes van a ser medidos? Para responder a ello se debe de tener claro cuál va a ser la **unidad de análisis**, que puede estar conformada por personas, un carácter positivo o negativo de las mismas, parámetros biológicos, objetos, una enfermedad..., que han de estar de acuerdo con los objetivos y el problema a investigar. El diseño de la muestra tiene grandes implicaciones metodológicas y es uno de los requerimientos técnicos destinados a elegir una representación adecuada de unidades de nuestra población objeto de estudio: El **muestreo** no es más que la elección de una parte de un todo que es la **población**.

EL OBJETIVO GENERAL de todo muestreo es llegar a conocer determinadas características de una población a partir de una selección de unidades de ésta, con el menor coste posible en dinero, tiempo y trabajo. Mediante las técnicas estadísticas, las leyes probabilísticas y los diseños muestrales basados en varios métodos de muestreo, podemos aproximarnos al conocimiento de estas características sin necesidad de tener que obtener la información exhaustiva de toda la población, garantizando la

representatividad y sabiendo que siempre **cometeremos un determinado error estadístico**, que se puede determinar de antemano en cada caso, por el hecho de tener una parte del todo.

Muestra

Una muestra estadística es una parte o subconjunto de unidades representativas de un conjunto llamado población o universo, seleccionadas de forma aleatoria, y que se somete a observación científica con el objetivo de obtener resultados válidos para el universo total investigado, dentro de unos límites de error y de probabilidad de que se pueden determinar en cada caso. Denotaremos al tamaño de la muestra mediante n.

En los estudios clínicos se pueden establecer las siguientes condiciones para definir las muestras:

1. Que comprendan parte del universo y no la totalidad.

2. Que el tamaño de la muestra sea estadísticamente proporcionado a la magnitud del universo.

3. Que se dé una ausencia de distorsión en la elección de los elementos de la muestra con el fin de evitar la introducción de sesgos que desvirtúen la representatividad.

4. Que sea posible poner a prueba hipótesis sustantivas de relaciones entre variables.

5. Que sea posible poner a prueba hipótesis de generalización, de la muestra en el universo, es decir, que sea representativa con un cierto grado de incertidumbre. En este sentido las muestras se dice que son probabilísticas y cualquier cálculo con los datos muestrales son estimaciones de características o parámetros poblacionales.

Los elementos a considerar antes de selección de la muestra aparecen recogidos en la tabla I.

Tabla I. Elementos a considerar en la selección de una muestra
1. Definir la población, tamaño y elementos que la componen
2. Determinar la unidad de observación, la unidad muestral y sus características
3. Recabar aquella información necesaria para hacer la selección de la muestra
4. Definir el tamaño de la muestra
5. Elegir el método para la selección de la muestra
6. Definir los procedimientos a seguir para la selección de la muestra

Mediante la **estadística descriptiva** alcanzamos el objetivo de describir la información estadística facilitando la tarea de análisis e interpretación de los datos mediante diversos cálculos (como los estadísticos descriptivos de la distribución de una variable y de las relaciones entre ellas), y cumplimos así la condición 4. Mediante la **estadística inferencial** fundamentamos los principios estadísticos que nos permiten alcanzar el objetivo de obtener generalizaciones estadísticas de la población a partir de la muestra, y cumplimos así con la condición 5. La estadística inferencial (o estadística inductiva) nos hablará de la significación de los cálculos en base al proceso de estimación de los parámetros poblacionales a partir de los estadísticos muestrales.

Universo o población

Universo y población son expresiones equivalentes para referirse al conjunto total de elementos que constituyen el ámbito de interés analítico y sobre el que queremos inferir las conclusiones de nuestro análisis. En particular se habla de población marco o universo finito, al conjunto preciso de unidades del que se extrae la muestra, y universo hipotético o población objetivo, el conjunto poblacional al que se pueden extrapolar los resultados. Denotaremos al **tamaño de la población** mediante N (Figura 1).

Figura 1. Relación de la muestra con la población a estudiar

Al hablar de poblaciones se establece la distinción entre una población finita y una infinita. Desde el punto de vista del muestreo, la distinción se basa en la importancia relativa que tiene el tamaño de la muestra n en relación al tamaño de población N. Si el tamaño de la muestra es muy pequeño respecto a la de la población (habitualmente se admite que represente menos del 5%) se suele considerar infinita la población. En cambio, si la muestra necesaria es considerable en relación a la población (por encima del 10% se suele considerar necesario, y entre un 5% y un 10% recomendable) se considera finita la población y se han de utilizar factores de corrección de población finita. Igualmente se considera que una población finita a toda población formada por menos de 100.000 unidades, e infinita a aquella que tiene 100.000 o más.

La **fracción de muestreo** indica simplemente el porcentaje que representa la muestra sobre la población. Cada uno de los elementos de una muestra o de la población se denomina unidad o individuo (ya sea una persona o no). El listado de todas las unidades de donde se extrae la muestra constituye la base o el marco de la muestra, también llamado técnicamente como espacio muestral.

El **error muestral** nos mide el grado de exactitud o de precisión con el que inferimos de la muestra a población. Este valor, como sugerimos anteriormente, vendrá determinado por la variabilidad del estadístico, es decir, que el error se cuantifica mediante las varianzas

del estadístico considerado en cada caso. Esta variabilidad se denominará error típico del estadístico, y se corresponde con un cálculo que depende de la varianza y del tamaño de la muestra, además de otra característica que es el nivel de confianza.

El error muestral es un aspecto clave de los estudios por muestreo que debemos considerar en dos momentos de la investigación: a) antes de proceder a obtener la muestra y de proceder a la recogida de información, en el que hay que determinar a priori qué nivel de error estadístico estamos dispuestos a asumir, junto a la decisión del tamaño de la muestra, y b) después de obtener la muestra, para determinar el error asociado a cada estimación o prueba estadística de hipótesis que podamos plantear.

¿Cuánto es un error muestral aceptable? Si lo expresamos en términos porcentuales, teniendo en cuenta en particular que los intervalos se construyen sumando y restando el valor del error, sería deseable alcanzar niveles del 2%. No siempre es posible alcanzar ese margen de error pues para ello se requiere un tamaño muestral relativamente elevado y, por ello, costes económicos no siempre asumibles.

Las fuentes de error sistemático se refieren a errores de medida y se pueden introducir en todo el proceso de investigación. Son fuentes de error sistemático:

- La elección de indicadores de los conceptos no adecuados.

- La inadecuada selección de la población y de las unidades. En particular, no disponer de listas completas, actualizadas y correctas de las unidades poblacionales a partir de las cuales efectuar la extracción aleatoria de las unidades que compondrán la muestra.

- La recogida de datos por encuesta se enfrenta con el problema del error de no respuesta, que se produce cuando no ha sido posible el contacto con la persona seleccionada (a veces resultado de considerar que es demasiado costoso llegar a

contactar) o bien se produce el rechazo de la persona a realizar la entrevista.

- Derivada del anterior se produce la incorrecta sustitución de los rechazos al realizar la entrevista.

- La pérdida de datos.

- La incorrecta consignación de las respuestas, la incorrecta codificación o registro no controlado de los datos en soporte informático.

- Las preguntas mal formuladas (sesgadas) o formuladas en un orden, con determinadas palabras, sobre determinadas características de la vida personal, etc.

- La ausencia de información por no respuesta a las preguntas del cuestionario.

- La inadecuada respuesta del entrevistado (por no recuerdo, incorrecta compresión, respuesta estereotipada o socialmente deseable, respuestas engañosas).

- Las perturbaciones o sesgos introducidos por el entrevistador/a, derivadas de una formación insuficiente, de unas condiciones de trabajo, de provocar un efecto de error no deseado en formular las preguntas, etc.

Alcanzar un bajo nivel de error muestral y de error sistemático son dos objetivos que se plantean en toda investigación por muestreo. Habrá que valorar ambos tipos de errores estableciendo un equilibrio óptimo en función de los recursos y del contexto de cada investigación.

Intervalo de confianza, nivel de confianza y nivel de significación

Las estimaciones de un parámetro de la población se construyen a partir de intervalos de confianza cuando se considera el nivel de error que se comete. Teniendo en cuenta el estadístico de la muestra y el error típico las estimaciones se expresan en términos de desviación

o variación que contempla el intervalo. Estas estimaciones por intervalo implican una afirmación probabilística que determina el nivel de confianza de una conclusión estadística. Así el **nivel de confianza** se define como la probabilidad de obtener el valor poblacional a partir de la muestra. Si decimos que el nivel de confianza es del 95% (o 0,95, en tanto por uno), estamos diciendo que para un tamaño muestral y desviación obtenidas, el parámetro poblacional se situará el 95% de las veces (o de las muestras) en un intervalo de valores que determina el valor del estadístico (o estimación puntual) sumándole y restándole el error muestral e. En el caso de la media (x), este es un intervalo definido por $x \pm e$, es decir, (x-e, x+e). Si el estadístico sigue una distribución muestral estadística normal, considerar un nivel de confianza del 95% implica afirmar que el 95% de los casos se sitúa entre los valores típicos –1,96 y 1,96, llamados límites de confianza. El i**ntervalo de confianza** es, por tanto, el conjunto de valores donde se encontrará el valor del parámetro poblacional con una probabilidad dada y un margen de error, y nos da una indicación de la exactitud de nuestra estimación. De forma complementaria se entiende por **nivel de significación** a la probabilidad de obtener un valor extremo del estadístico que estima el parámetro poblacional. Si el nivel de confianza es del 95%, el de significación es del 5% (o 0,05, en tanto por uno).

Recuerde que si desea un margen de error más pequeño, debe tener un tamaño de muestra más grande para la misma población. Cuanto más alto desee que sea el nivel de confianza, más grande tendrá que ser el tamaño de la muestra.

Error sistemático

Tan importante como el error estadístico es el error sistemático, es decir, toda aquella fuente de error no estadística que puede derivarse resultado de cualquier actuación incorrecta a lo largo de todo el proceso de investigación, en particular, por ejemplo, como resultado de un error de selección de la muestra o como resultado de una inadecuada medida.

Tamaño de la muestra

En un estudio resulta imposible estudiar a la población entera. Cada tipo de estudio tiene un tamaño muestral idóneo que permite comprobar lo que se pretende con seguridad aceptable y el mínimo esfuerzo posible.

Si se trata de un estudio transversal hemos de tomar en consideración que el objetivo es estimar el dato de población a partir de la muestra. Para ello tendremos que:

- Elegir una variable para el cálculo
- Decidir un nivel de confianza
- Decidir el máximo error aceptable
- Considerar el tipo de muestreo

Formulas para el cálculo del tamaño muestral: para el cálculo en cada tipo de estudio existe una fórmula estadística apropiada. Se basan en el error estándar, que midel el intervalo de confianza de cada parámetro que se analiza (media aritmética, porcentaje, diferencia de medias etc. La precisión estadística aumenta (el error estándar disminuye) cuando el tamaño de la muestra crece.

- Cuando se conoce N

$$n = \frac{N * Z^2 * \sigma^2}{(N-1) * E^2 + Z^2 * \sigma^2}$$

- Cuando no se conoce N

$$n = \frac{Z^2 * \sigma^2}{E^2}$$

Donde "n" es el tamaño de la muestra; z el nivel de confianza, E el margen de error (porcentaje expresado en decimales), $\sigma 2$ la varianza poblacional, y N el tamaño de la población.

Para encontrar la puntuación z adecuada, consulta la tabla II.

Tabla II. Nivel de confianza deseado y puntuación z

Nivel de confianza deseado	Puntuación z
80 %	1.28
85 %	1.44
90 %	1.65
95 %	1.96
99 %	2.58

Actualmente se dispone de programas o calculadoras en línea que se pueden descargar de forma gratuita, que se manejan de manera sencilla, para establecer el tamaño de la muestra que precisamos para el estudio. Calculador en línea: www.openepi.com. Software gratuitos: Epidat 3.1 o 4.2 https://www.sergas.es/Saude-publica/EPIDAT-4-2?idioma=es, https://www.sergas.es/Saude-publica/Epidat-3-1-descargar-Epidat-3-1-(espanol) y G Power: https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower

Tipos de muestreo

El diseño de una muestra puede seguir estrategias distintas que identifican diferentes tipos de muestreo. Una primera distinción fundamental es la que los clasifica en función de si son probabilísticos o no. Nuestro mayor interés en este texto es con relación a los primeros. Daremos cuenta de ellos en los apartados siguientes, pero presentamos simplemente la relación de tipos de muestreo.

a) **Muestreo probabilístico.** Aquel muestreo en que, de forma estricta, todas las unidades de la población tienen una probabilidad conocida de ser incluidas en la muestra, y, por lo tanto, también se conoce la probabilidad de obtener cada una de las muestras mediante un procedimiento de aleatorización. En esta categoría se encuentran:

- Muestreo aleatorio simple
- Muestreo sistemático
- Muestreo estratificado
- Muestreo por conglomerados
- Muestreo polietápico

En el muestreo probabilístico, además, se puede considerar una distinción cuando los diferentes elementos de la muestra o bien son idénticos o similares y tienen la misma probabilidad de ser elegidas, o bien esta probabilidad es diferente. Así se diferencia entre muestreo con probabilidades iguales y muestreo con probabilidades desiguales. También el muestreo tiene implicaciones en función de si se trata de muestreo con reposición o sin reposición (o reemplazamiento).

b) **Muestreo no probabilístico**. En este caso no se conocen las probabilidades de cada unidad de muestreo de pertenecer a la muestra. Cabe contemplar en esta categoría:

- Muestreo por cuotas
- Muestreo casual o incidental
- Muestreo de conveniencia
- Muestreo intencional o razonado
- Muestreo de bola de nieve

Cada tipo de muestreo tiene sus indicaciones y sus limitaciones. Y el modo de llevarlo a cabo exige un estudio particular para uno de ellos. Evidentemente, el análisis y la puesta en práctica de cada uno mediante ejemplos exigiría un capítulo específico y un estudio prolon-

gado. Debe saberse que para facilitar la realización del muestreo existen diferentes herramientas informáticas. La más sencilla es la hoja de cálculo Excel, que genera números aleatorios que podemos asignar a cada sujeto de la población y luego permite ordenarlos hasta alcanzar el tamaño muestral deseado. Una alternativa mejor es utilizar el programa Epidat, software gratuito que puede ser descar4gado libremente. Para profundizar en esta materia recomendamos la lectura del texto elaborado por el profesor Carlos Ochoa y referido en la bibliografía.

Tareas implicadas en el diseño de una muestra

En el diseño de una muestra cabe contemplar las siguientes tareas relevantes: 1) Ante todo, la definición de los objetivos de investigación, la construcción de un modelo de análisis y determinación de los recursos disponibles. 2) La precisión de las variables y parámetros poblacionales a medir o que expresan el centro de interés de la investigación. 3) La delimitación de la población objetivo y de la población marco. Definición de las unidades de muestreo y medidas. Delimitación espacial y temporal. 4) Constituir la base de la muestra o marco de muestreo, cuando sea posible. 5) Elección del tipo de muestreo. 6) Determinar el tamaño de la muestra, con un nivel de confianza y error muestral dados. 7) Extracción aleatoria de la muestra. 8) Organización del desarrollo del trabajo de campo, recogida de información y seguimiento. 9) Validación y ajuste de la muestra.

Definir las variables

El concepto de variable está en el núcleo de la investigación cuantitativa. Como definición general, una variable es un rasgo mensurable de un caso concreto: un caso es una "cosa" concreta o una unidad que muestra este rasgo mensurable. Tal como expresa el término, los valores (o resultados) de este rasgo pueden variar entre los casos, pero cada caso solo puede ofrecer un valor para un rasgo concreto.

Estrechamente ligado al concepto de variable está el concepto de medición. Medición es el proceso de asignar un valor específico de

una variable a un caso concreto, usando en ello criterios predefinidos. La medición, por tanto, significa colocar un objeto o persona concreta (un "caso") en una categoría concreta.

Tipos de variables

Los tipos de variables vienen determinados por el dato que representa. Por ejemplo, el peso es una variable cuantitativa cuando se expresa en números como gramos o kilogramos de un objeto. Mientras que si se presenta en términos de "pesado" o "ligero", sería una variable cualitativa, porque presenta una cualidad.

Una **variable cuantitativa discreta** solamente puede tomar valores integrales, es decir 1, 2 o 555, pero no 1.5 o 2.25. Ejemplos de este tipo de variables son:

- El número de veces que algo sucede
- El número de veces que alguien asume un determinado comportamiento
- La cantidad de personas o seres en un grupo
- La cantidad de objetos en un lugar

La **variable cualitativa dicotómica** es un dato no numérico que presenta una cualidad, propiedad o condición observable, que nada más presenta dos valores. Por ejemplo:

- El veredicto de un jurado: "culpable" o "no culpable".
- El sexo: "masculino" o "femenino".
- El resultado de un examen de antígeno: "positivo" o "negativo".
- Presencia de una condición: "presente" o "ausente".
- El tipo de hospital: "público" o "privado".

Variable cualitativa categórica o nominal es la variable no numérica que presenta tres o más categorías. Por ejemplo: deportes olímpicos: "natación", "voleibol", "atletismo", "esgrima" o "gimnasia"; estados de la materia: "sólido", "líquido" o "gaseoso".

Variables ordinales, en las que los valores pueden ordenarse, de menor a mayor, de más importante a menos importante, de primero a último, etc. Este tipo de variable la observamos en: clase social: "clase baja", "clase media" o "clase alta"; competencia en un idioma: "básico", "intermedio" o "avanzado".

Las variables también se pueden clasificar según que se presenten sin necesidad de otra (**variable independiente**) o que sea consecuencia de otra (**variable dependiente**) (Figura 2). Por lo general, los estudios científicos se enfocan en examinar los efectos de una variable independiente. En un estudio se analizó el impacto de cinco intensidades de un campo magnético sobre plantas de cebada. En este caso, la variable independiente fue la intensidad del campo magnético.

Figura 2. Relación entre variable independiente y variable dependiente

La variable dependiente es la medida del efecto de la variable independiente. La forma más fácil de identificar una variable dependiente es detectando el efecto o la consecuencia de algo, es decir, la variable independiente que es la causa. Mientras la variable independiente se manipula o fija, la variable dependiente se mide o registra.

Describir y analizar los datos

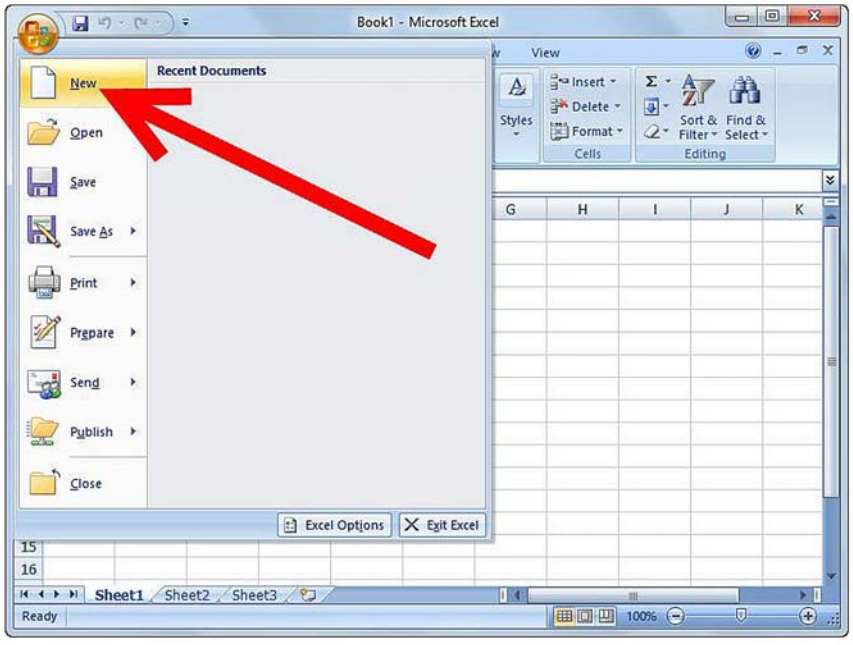
La descripción y análisis de datos es el proceso de examinar, limpiar, transformar y modelar un conjunto de datos con el objetivo de descubrir información útil, extraer conocimientos y tomar decisiones informadas. Implica el uso de técnicas y herramientas estadísticas, matemáticas y de visualización para identificar patrones, tendencias y relaciones en conjuntos de datos. El análisis de datos permite revelar insights, responder preguntas y resolver problemas, ayudando a las or-

ganizaciones y personas a comprender mejor el mundo que les rodea, optimizar procesos y tomar acciones basadas en evidencia.

Bases informáticas y programas básicos

Las bases de datos son conjuntos de variables (datos u observaciones) organizados bajo una estructura común. Cada elemento de la estructura se denomina Campo de la base datos, y al conjunto de campos de un unidad de estudio lo denominamos registro.

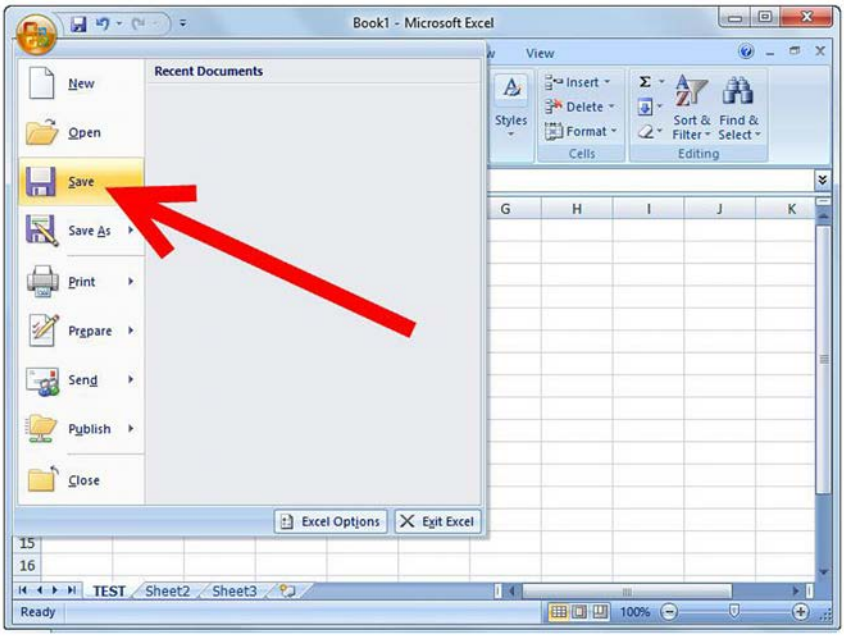
Los datos de la base se almacenan en tablas de filas y columnas. Las casillas de la fila superior indican los nombres de las variables de su columna correspondiente, mientras que cada fila contiene los distintos campos de cada registro. Cada casilla de la tabla es un elemento de información.



Entre las alternativas para introducir datos tenemos las hojas de cálculo (Excel, por ejemplo), gestores de bases de datos (Access, por ejemplo), paquetes estadísticos (SPSS, R) y aplicaciones on-line (Gooi-

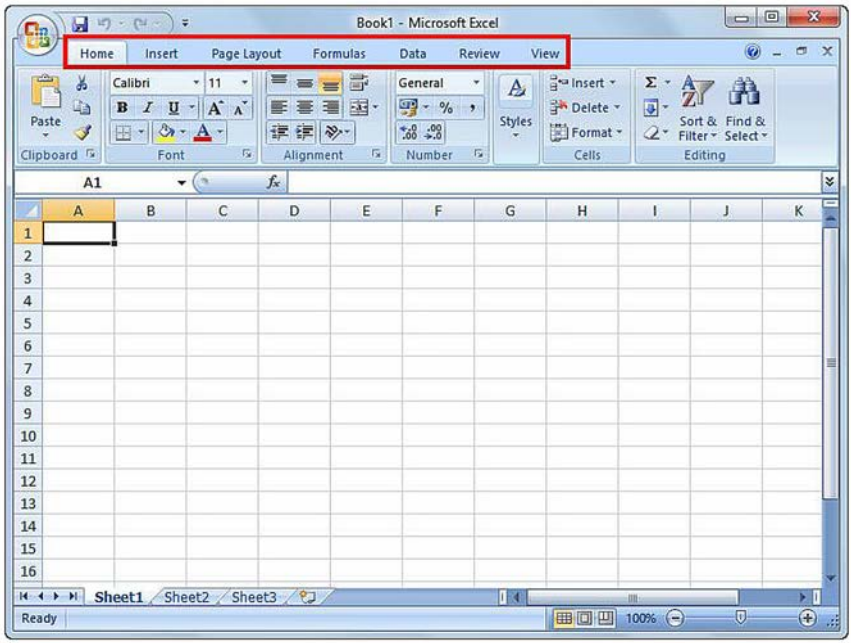
gle Forms, por ejemplo). En este capítulo nos detendremos en aquellas alternativas de uso más común.

Hojas de cálculo



Excel es un programa de hoja de cálculo de gran alcance fabricado por Microsoft Office. Con él puedes crear y dar formato a una hoja de cálculo y a un libro (archivo que contiene una o más hojas de cálculo), construir modelos para el análisis de datos, escribir fórmulas, realizar muchos cálculos, y presentar gráficos profesionales. Es posible que Excel tenga un acceso directo en el escritorio, o es probable que tengas que hacer clic en "Inicio" y luego en "Programas" para localizar el ícono de Excel.

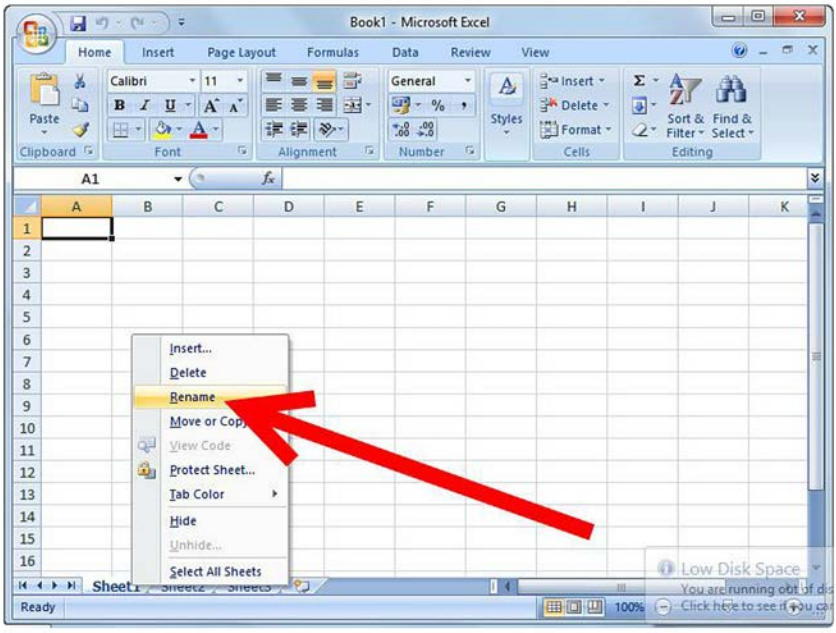
Ha de tenerse en cuenta que lo que se muestra en el presente apartado son apenas instrucciones básicas para utilizar Excel. Para



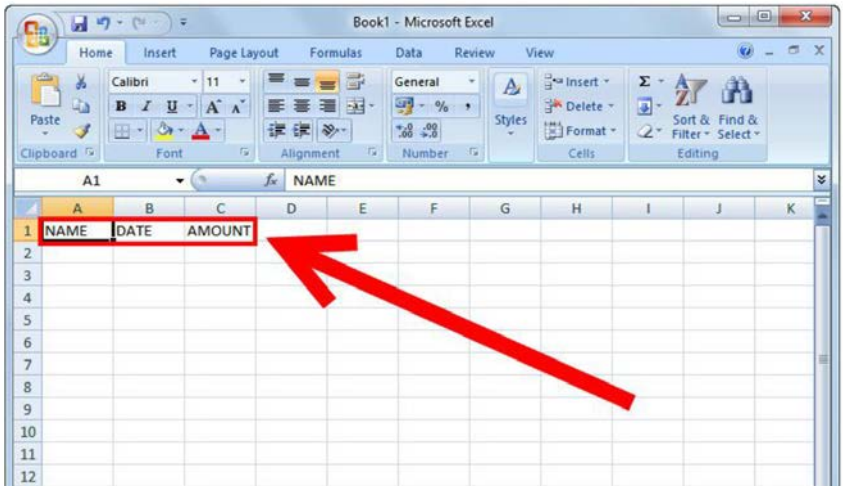
aprender a manejar correctamente cualquier programa de manejo de datos es necesario contar con un buen manual de referencia y realizar algún curso de formación que nos introduzca en las destrezas que se precisan. Como alternativa, poder contar con un experto cercano que nos resuelva las dudas y llame la atención sobre sus aplicaciones y los errores más comunes. En lo que se refiere a nuestro trabajo de investigación las utilidades de uso más común son:

1. Configurar una hoja de cálculo
2. Ingresar y gestionar datos en Excel
3. Realizar cálculos básicos

Para Iniciar un nuevo libro (un archivo de Excel): hacer clic en "Archivo" y luego en "Nuevo". En "Plantillas disponibles", haz clic en "Libro en blanco" y luego en "Crear". A continuación, se abrirá un libro en blanco.



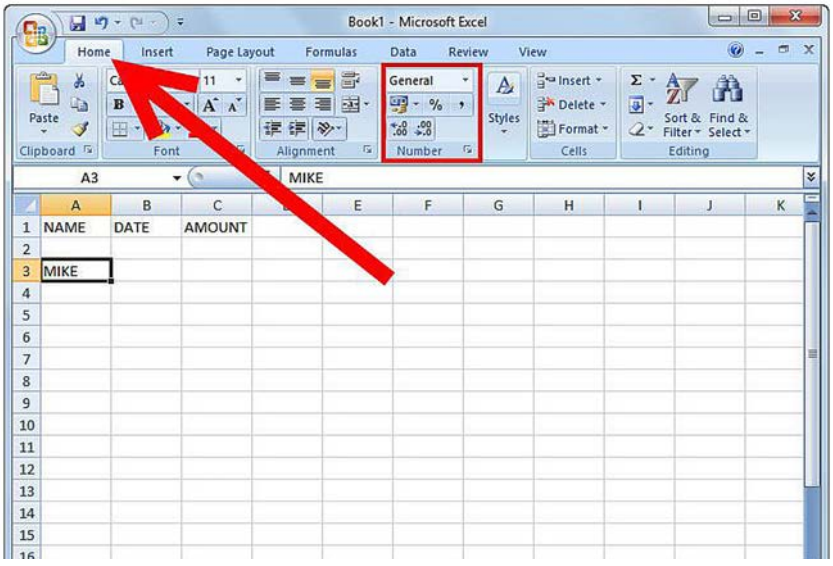
Para guardar el libro hacer clic en el botón de Office (o en la pestaña Archivo si cuentas con una versión anterior de Excel) y seleccionar "Guardar como". Escoger una ubicación en el ordenador guardar



el archivo (por ejemplo, la carpeta "Mis documentos"), poner el nombre del libro en el cuadro "Nombre de archivo", y asegúrese de que el tipo de archivo esté configurado como "Libro de Excel".

Es necesario familiarizarse con las pestañas de la cinta en la parte superior del libro. Estas son "Archivo", "Inicio", "Insertar", "Diseño de página", "Fórmulas", "Datos", "Revisar" y "Vista". Es importante conocer los términos más comunes en tecnología para poder utilizar con facilidad y comprender mejor las guías de este tipo de programas. Una fila es una sección que cruza la pantalla de derecha a izquierda. Las filas están indicadas por números a lo largo del lado izquierdo de la pantalla. Una columna es un conjunto de datos que va desde la parte superior hasta la parte inferior de la hoja de cálculo, y se identifica con una letra en la parte superior de la hoja. Una celda es cualquier cuadrado individual de la hoja de cálculo dentro del cual se colocan los datos.

Preparar la hoja de cálculo para almacenar los datos. Cada libro de Excel tiene 3 hojas de trabajo predeterminadas. La "Hoja 1" se

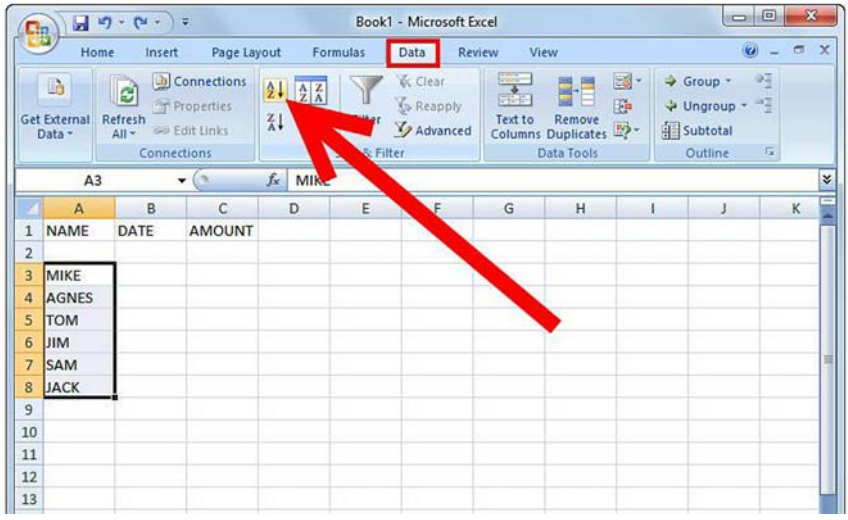


abre de manera predeterminada y puedes encontrar su pestaña en la parte inferior de la ventana.

Cambiar el nombre de una hoja haciendo clic derecho sobre la pestaña "Hoja 1": Seleccionar la opción "Cambiar nombre" y escribe el nuevo nombre para la hoja. Si es necesario, agregar más hojas haciendo clic en el botón a la derecha de "Hoja 3", que muestra una hoja de papel con una estrella en la esquina.

En la fila superior de la hoja, escribir un título en cada celda para identificar aquello que se va a colocar en cada columna. Por ejemplo, se puede escribir "Nombre", "Fecha" y "Cantidad". Las filas debajo de estos títulos son para colocar los datos. Guardar los avances con frecuencia. A medida que se introducen los datos, debe guardarse el trabajo con frecuencia haciendo clic en el símbolo de disquete en la parte superior izquierda de la pantalla, o haciendo clic en el botón de Office y seleccionando la opción "Guardar". Como alternativa, puede mantenerse presionada la tecla "Ctrl" en tu teclado mientras presionas la tecla "G".

Comenzar a **ingresar los datos en las celdas** de la hoja de cálculo. Es posible decidir comenzar con una o dos columnas o filas para



practicar antes de ingresar todos los datos en la hoja. Hacer clic en una celda y escribir los datos. Para editar los datos después de ingresarlos en la celda, hacer doble clic en la celda o editar los datos en la barra de edición en la parte superior de la hoja de cálculo (la que se encuentra justo por encima de las letras de las columnas).

Aprender a **darle formato a las celdas**. Se utiliza el formato "General" de manera predeterminada, pero pueden cambiarse las configuraciones de cada celda, fila o columna. Este formato se puede cambiar a uno preestablecido (como "Número", "Fecha", "Hora", o "Moneda") seleccionando la flecha desplegable junto a "General" en la pestaña "Inicio". Además, puede cambiarse el tipo de letra y estilo, así como la alineación de los números o el texto utilizando las secciones de "Fuente" y "Alineación" de la pestaña "Inicio". Cambiar el formato de una columna entera, seleccionando la letra en la parte superior y luego realizando los cambios. Cambiar el formato de una fila entera seleccionando el número en el lado izquierdo de la pantalla y luego realizando los cambios.

Ingresar los datos. Para a agregar tos los datos a la hoja de cálculo. Presionar la tecla "Enter" para pasar a la siguiente celda (debajo de la celda actual). Presionar la tecla "Tab" para pasar a la celda de la derecha, o utiliza alguna de las teclas de dirección para cambiar de celda. Recordar que siempre debes guardar tus avances.

Para **ordenar los datos**, seleccionar los que deseas ordenar. Si se desea, pueden seleccionarse columnas individuales o columnas múltiples e incluir títulos de texto. Asegurarse de seleccionar varias columnas si deseas mantener las filas de datos juntas. Si se ordena una sola columna, el orden de la columna cambiará, pero dejará las columnas adyacentes sin ordenar. Selecciona la pestaña "Datos" y haz clic en "Ordenar". Aparecerá un cuadro de diálogo de ordenación. Seleccionar la columna que desea ordenar en la lista "Ordenar por". Si se ingresan títulos en la fila superior, los títulos de tus columnas aparecerán en el cuadro "Ordenar por". Selecciona "Valores", "Color de celda", "Color de fuente" o "Ícono de celda". Si se ingresa un texto, es probable que se desee seleccionar "Ordenar por Valores". Seleccionar el

orden en que deseas aplicar la operación de ordenación. Este puede ser ascendente o descendente ("de la A a la Z" o "de la Z a la A" para textos, o "de mayor a menor" o "de menor a mayor" para números).

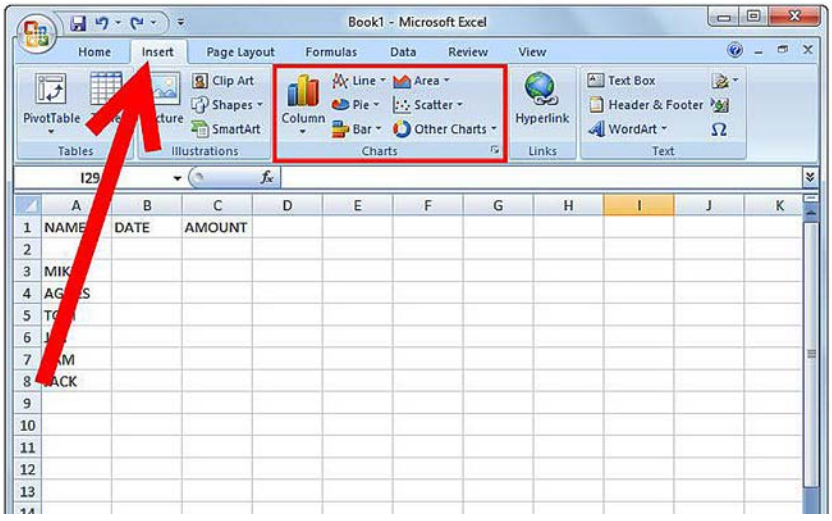
Filtrar los datos: seleccionar los datos que desea filtrar resaltando una o varias columnas. Seleccionar la pestaña "Datos" y hacer clic en "Filtro" (el ícono de embudo) en la sección "Ordenar y filtrar". Las flechas aparecerán en la parte superior de cada columna. Hacer clic en la flecha para ver la lista de opciones en el encabezado de la columna. Seleccionar los valores que se desea utilizar y haz clic en "Aceptar" para ver los resultados. El resto de los datos se ocultará para que se pueda ver solo los datos filtrados. Restaurar el resto de los datos seleccionando la opción "Borrar" (el ícono que muestra un embudo con una "X" roja junto a este) en la sección "Ordenar y filtrar" de la pestaña "Datos".

Buscar textos específicos en tu libro: hacer clic en el ícono "Buscar y seleccionar" (binoculares) en la pestaña "Inicio". Hacer clic en "Buscar" y escribir el texto que se está buscando. Seleccionar "Buscar todos" y aparecerá una lista con todas las instancias de ese texto en la hoja de cálculo. Nota: para buscar en todo el libro, selecciona el botón "Opciones" en la ventana emergente "Buscar y reemplazar" y en la opción "Dentro de", cambia "Hoja" por "Libro", y luego haz clic en "Buscar todos".

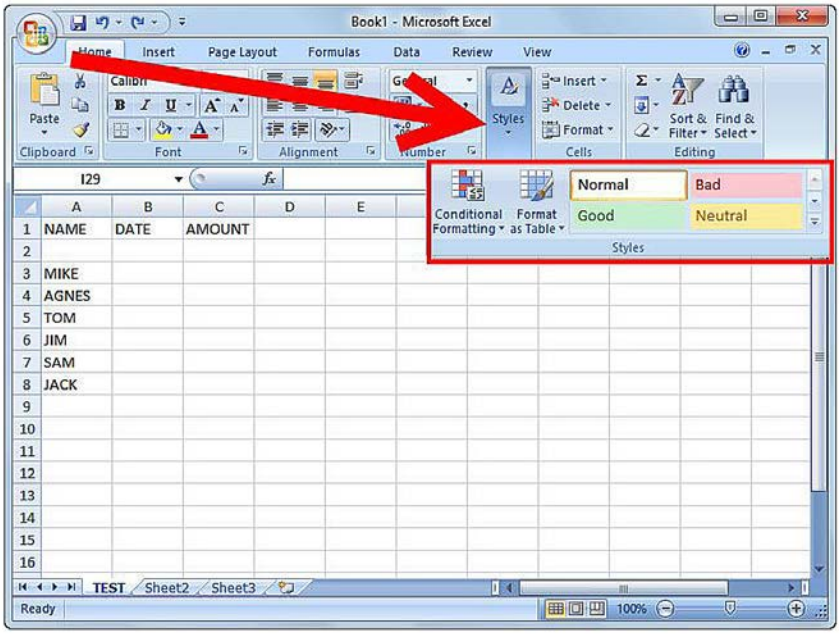
Para **imprimir la hoja de trabajo** haciendo clic en "Archivo" y luego en "Imprimir", o manteniendo presionada la tecla control en tu teclado mientras presionas la tecla "P" (comando de teclado: "Ctrl + P"). Puede obtenerse una vista preliminar del archivo haciendo clic en el botón de Office y luego en "Imprimir". Luego, se puede seleccionar la opción "Vista preliminar". A partir de allí, puede seleccionarse el ícono de impresión en la parte superior izquierda de la pantalla. También se puede cambiarse la configuración, el tamaño, los márgenes y otras opciones de página accediendo al menú "Configurar página" en la pestaña "Diseño de página". Seleccionar la flecha pequeña en la esquina de la sección "Configurar página" para expandir el menú. Ajustar toda la hoja

de cálculo para que encaje en una sola página impresa accediendo a la pestaña "Diseño de página". Luego, hacer clic en la flecha pequeña al lado de "Ajustar a". Debajo de "Ajuste de escala" en la pestaña "Página", seleccionar "Ajustar a" y cambia la configuración a "1 página de ancho por 1 de alto". Luego, hacer clic en "Aceptar". Para imprimir solo una parte, haz clic sobre la hoja de cálculo y seleccionar el rango de datos que se desea imprimir haciendo clic, manteniendo pulsado el botón del ratón y arrastrándolo para cubrir la selección que se desea. Luego, hacer clic en el botón de Office, seleccionar "Imprimir" y selecciona "Imprimir Selección" en "Configuración". Luego, hacer clic en "Aceptar". Utilizar el cuadro desplegable debajo de "Impresora" para ver las impresoras que se encuentran instaladas en el equipo y selecciona la impresora que deseas utilizar.

Utilizar la función de sumatoria para realizar una suma básica en la columna. Hacer clic en la celda vacía debajo de una columna de números que deseas sumar. Hacer clic en el símbolo "Autosuma" al lado derecho de la cinta de la pestaña "Inicio" (que parece una "M" de costado). Hacer clic una segunda vez y la celda previamente vacía ahora mostrará el total de la columna de números.



Utilizar un signo igual (=) para comenzar una fórmula. Si se va a ingresar una fórmula a mano (en lugar de utilizar el botón de sumatoria), debe comenzarse la fórmula con un signo igual. Coloca primero el signo igual en la celda en la que deseas que aparezca la respuesta.



Sumar una columna entera de números: Puede realizarse esto utilizando la Autosuma, pero también lo puedes realizar escribiendo la fórmula tú mismo. Digitar el signo "=" en una celda vacía (en la que se desea que aparezca la suma), y luego, escribe "SUM". Digitar un paréntesis de apertura "(", luego, la letra de la columna y el número del rango superiores (primeros) que deseas sumar. A continuación, digita el signo de dos puntos. Escribir la letra de la columna y el número del rango inferiores (últimos) que se desea sumar. A continuación, cerrar el paréntesis para encerrar las letras y los números. Por ejemplo, la fórmula debe quedar de la siguiente manera: "=SUM(B5:B9)". Presionar "Enter" para que tu columna de números se sume.

Sumar números utilizando una fórmula propia: si desea sumar números que no están alineados en una columna, puede crearse una fórmula propia para sumar. Hacer clic en la celda vacía en la que quiere que aparezca la suma. Teclear el signo "=". Luego, hacer un clic en la primera celda que deseas sumar o escribir la letra de la columna y el número de fila correspondiente a tu primer número (por ejemplo, "B2"). Digitar el signo "+" y luego hacer clic en la siguiente celda que desea sumar o escribir la letra de la columna y el número de fila correspondiente al segundo número. Por ejemplo "=B2+C5". Repitir hasta que se hayan seleccionado todos los números que desea sumar. Presiona "Enter" y aparecerá la respuesta. Utiliza el mismo método para restar, cambiando el signo "+" por "-".

Seleccionar los **datos que se desea que aparezcan en un gráfico**: ubicar la pestaña "Insertar" y la sección "Gráficos". Hacer clic en el tipo y subtipo de gráfico que se desea utilizar. Crear una tabla seleccionando un rango de celdas. Todas las celdas pueden contener los datos, o algunas pueden estar vacías. Buscar la pestaña "Inicio" y la sección "Estilos". Luego, hacer clic en "Dar formato como tabla". Seleccionar el estilo de la tabla a partir de las muchas opciones que aparecen.

Crear una tabla seleccionando un rango de celdas: todas las celdas pueden contener los datos, o algunas pueden estar vacías. Buscar la pestaña "Inicio" y la sección "Estilos". Luego, hacer clic en "Dar formato como tabla". Seleccionar el estilo de la tabla a partir de las muchas opciones que aparecen.

Sombrear o colocar bordes a las celdas: seleccionar el rango de celdas al que desea aplicarle esto. Ir a la pestaña "Inicio" y luego a la sección "Fuente". Hacer clic en "Fuente" y en la pestaña "Bordes". En el cuadro de diálogo, seleccionar el estilo requerido. Para sombrear celdas, ir a la pestaña "Inicio" y a la sección "Fuente". Hacer clic en "Fuente" y en la pestaña "Relleno", y seleccionar el color de sombreado que desees.

En diferentes partes pudo haber leído que Excel es algo limitado, o incluso deficiente. Sin embargo, para determinados análisis estadísticos es el programa de hojas de cálculo más utilizado a nivel mundial. Las cosas que Excel hace muy bien son:

- Nos permite introducir, ordenar y filtrar datos de manera rápida y eficaz.
- Permite crear diagramas y gráficos de forma bastante rápida y eficaz.
- Hace operaciones aritméticas básicas, como sumas/restas, multiplicaciones/divisiones, cálculos de frecuencia absoluta y relativa.
- Hasta cierto punto, calcula medidas de centralización y de dispersión.
- Al formar parte del entorno Microsoft Office muchos usuarios están familiarizados con su apariencia y su uso.

Programas alternativos y necesarios al avanzar en experiencia investigadora

Si lo que se necesita es una forma rápida y sencilla de analizar datos, o si se necesita tener una idea general del aspecto de los datos, Excel cumplirá su misión. Peros si lo principal es la precisión o si el trabajo requiere métodos estadísticos más complejos debemos dar el paso hacia dos programas informáticos propiamente dichos y muy usados: SPSS y R.

SPSS es el acrónimo de Statistical Package for Social Sciences. Es una herramienta muy popular. No es una hoja de cálculo, aunque su interfaz principal parezca una tabla compuesta de celdas. Para operar con este programa es preciso especificar cada variable individual: no solamente su nombre, sino su tipo (si son numéricas, textuales o de cualquier otro tipo) y su nivel de medición. Podemos especificar también los valores "ausentes", pidiéndole al programa qué hacer cuando

careemos de datos para un caso/participante concreto, una cosa que no es fácil de hacer con Excel. Es muy importante que las variables estén correctamente especificadas: si entra basura, sale basura (o GIGO en sus iniciales en inglés: Garbage in, garbage out). Aparte de media, mediana, moda, desviación típica, varianza, rangos, tables de frecuencia, estadísticos descriptivos, frecuencias, permite realizar análisis de correlación, valores de significación, análisis de regresión, prueba t y ANOVA, MANOVA, etc.

Es cada vez más frecuente el empleo habitual del programa informático llamado R: "R" es un entorno informático para la computación y gráficos estadísticos. No es un programa estadístico como SPSS, o una hoja de cálculo como Excel, sino que es en realidad un lenguaje de programación. Esto significa que ofrece una herramienta potentísima que permite hacer pruebas estadísticas complejas con un grado alto de precisión, pero no mediante interfaz amable y los fáciles toque de ratón de SPSS y Excel. El libro del profesor Carlos Redondo sirve como referencia para introducirse en el uso y explotación del programa y de su entorno de programación.

Fusión, división y reestructuración

La tarea de transformación de los datos está destinada a adaptar los datos a las necesidades del análisis donde se requiere modificarlos, para realizar correcciones y cambios en la información existente inicialmente, ya sea en relación a las variables de un archivo de datos o en relación al tratamiento de varios de ellos, o para generar nuevas variables basadas en las existentes: agrupaciones, tipologías, índices, etc. Como en el apartado anterior presentaremos en dos subapartados distintos los procedimientos de transformación para SPSS y R.

Programa SPSS.

La fusión o unión de archivos da lugar a dos alternativas:

- Añadir variables: se fusiona el archivo de datos activo con otro que contiene los mismos casos pero variables diferentes.

- Añadir casos: se fusiona el archivo de datos activo con otro que contiene las mismas variables pero casos diferentes.

Una necesidad habitual en el tratamiento de los datos de un fichero es segmentarlo, es decir, **dividirlo en grupos de individuos** según los valores de una o más variables de agrupación para realizar un mismo tipo de análisis que se repetirá dentro de cada grupo. Para poder realizar a la segmentación correctamente será necesario ordenar previamente el archivo. El SPSS nos ofrece dos formas diferentes de segmentar el archivo: 1) Comparar los grupos: los grupos se presentan juntos para poder compararlos en una sola tabla o con gráficos individuales que se presentan juntos. Y 2) Organizar los resultados por grupos: los resultados de cada procedimiento se muestran por separado para cada grupo. El comando de segmentación es SPLIT FILE (menú Datos / Segmentar archivo). Esta opción tiene diversas aplicaciones, pero una de ellas podría ser la de elaborar el anexo estadístico con numerosas tablas y gráficos que queremos repetir, por ejemplo, para cada territorio del estudio por separado. Aquí de nuevo es importante recordar que una vez hayamos realizado el análisis deseado es necesario deshacer la segmentación para volver a trabajar con el archivo completo, como una sola muestra. Para ello volvemos al menú y marcamos Analizar todos los casos.

La **agregación de casos** tiene múltiples usos en el tratamiento de matrices de datos, también cuando se relacionan diversas bases. Definidos los cálculos podemos optar por tres alternativas: 1) Añadir variables agregadas al conjunto de datos activo. Las nuevas variables calculadas de grupo son un atributo de cada unidad de la base de datos original por lo que cada caso con los mismos valores de segmentación recibe los mismos valores para las nuevas variables agregadas. 2) Crear un nuevo conjunto de datos que contenga únicamente las variables agregadas. Se crea un nuevo conjunto de datos en la sesión actual con las variables de agregación y agrega las unidades. Y 3) Escribir un nuevo

archivo de datos que contenga sólo las variables agregadas. Es el caso anterior pero guardando los datos agregados en un archivo de datos externo que hay que detallar.

La estructura simple de una matriz de datos de casos por variables suele ser la habitual para el análisis de datos, no obstante, la estructura inicial de una base de datos puede ser compleja. Una estructura donde la información de una variable está en más de una columna o la información de un caso en más de una fila introduce una complejidad de organización de la información y la necesidad de reestructurar el archivo para pasar los casos a variables o las variables a casos. Por ejemplo, si tenemos una matriz con 3 individuos y las condiciones de empleo se refieren a dos momentos en el tiempo: empleo inicial y empleo actual, la información puede estar dispuesta por filas donde cada individuo tiene doble información de sus condiciones de empleo, la inicial y la actual.

Pasa **fusionar** es muy conveniente disponer de una variable clave que identifique a cada unidad en cada uno de los archivos a unir, de esta forma se irá emparejando la información a partir del control de la coincidencia del mismo caso. El tipo de fusión que haremos implicará que ambos archivos proporcionan casos individuales en los dos archivos.

Transformación de los datos con R

Dado que incluye algunos procedimientos destinados al tratamiento de ficheros, ya sea en su interior, ya sea para combinarlo con otros, y de transformación para la creación de variables, distinguiremos dos tipos de procedimientos de gestión y transformación de archivos: los destinados al tratamiento de datos en el interior de un fichero y al tratamiento de datos entre ficheros que se relacionan.

La **fusión o unión de archivos** da lugar a dos alternativas:

- Añadir variables. Se fusiona el archivo de datos activo con otro que contiene los mismos casos pero variables diferentes.

- Añadir casos. Se fusiona el archivo de datos activo con otro que contiene las mismas variables pero casos diferentes.

Depuración de la base de datos

La depuración de datos, también llamada limpieza de datos o scrubbing, es el proceso de modificación o eliminación de datos en una base que es incorrecta, está incompleta, tiene un formato incorrecto o está duplicada. Elimina errores de tipo lógico que pueden presentarse en los datos ya grabados y aportar un procedimiento sistemático original para detectarlos y corregirlos. Para ello se ha programado un conjunto de macros SPSS que permiten detectar todos estos errores, generar de forma totalmente automática un informe de incidencias para corregirlos y ofrecer una estadística final de errores. Los algoritmos de estos macros son fácilmente transportables a SAS o a otros sistemas. El procedimiento propuesto consiste en crear un archivo de sintaxis con un conjunto de llamadas a macros que realizan el proceso de acuerdo con las siguientes fases: 1) lectura de la tabla con los datos originales grabados; 2) depurar el identificador para garantizar que cada registro está unívocamente identificado y se adecua a las formas normales de integridad referencial de la teoría relacional; 3) corregir las incidencias detectadas en el identificador; 4) incorporar las variables de referencia de otras tablas que sean necesarias para depurar la tabla de datos; 5) depurar las variables de salto; 6) corregir las incidencias detectadas en las variables de salto; 7) depurar el resto de variables del estudio, detectando las incidencias que sean consecuencia de inconsistencias y los valores desconocidos; 8) corregir las incidencias detectadas, introduciendo el valor correcto o valor desconocido si no se conoce el valor correcto; y 9) generar una estadística de los errores detectados y de los valores desconocidos presentes en la matriz de datos depurados.

El proceso comporta realizar de forma iterativa las fases de chequeo y corrección hasta que las únicas incidencias detectadas sean valores desconocidos no recuperables. Asimismo, el proceso incorpora

un historial de cambios que permita conocer todas las modificaciones efectuadas sobre los datos originales.

Bibliografía

REDONDO FIGUERO C. El programa R, herramienta clave en investigación. Editorial Universidad de Cantabria, Textos Universitarios Nº 23. Santander 2016. ISBN 978-84-8102-797-6

OCHOA SANGRADOR C. Diseño y análisis en investigación. IMC, Madrid 2019. ISBN 978-84-7867-685-9.

FIELD AP. Discovering Statistics Using SPSS (ans Sex, Drugs and Rock ´N´ Roll). London, Sage 2009.

PARDO A, RUIZ MA. Análisis de datos con SPSS 13. Madrid 2005, McGraw-Hill.

PARDO A, RUIZ MA. Gestión de datos con SPSS Statistics. Síntesis, Madrid 2009.

RIAL A, VARELA J, ROJAS AJ. Depuración y análisis preliminares de datos en SPSS.RA-MA, Madrid 2001.

ARGIMÓN JM, JIMÉNEZ J. Estudios experimentales I: el ensayo clínico aleatorio. En: Métodos de investigación clínica y epidemiológica. Elsevier España SA. Madrid, 2004: 33-48.

La prueba estadística de la información

Javier Ochoa Brezmes.

Análisis básico de datos

EL OBJETIVO DE la investigación clínica es obtener información de la población para responder a preguntas clínicas. Como hemos visto en capítulos previos recabamos datos de muestras de pacientes que posteriormente debemos analizar. Para que las estimaciones realizadas a partir de los datos sean válidas y aplicables a la población las muestras deben ser representativas. El proceso comienza con una exploración descriptiva de las variables recogidas, lo que constituye el análisis básico, sobre el que se sustentará el análisis avanzado, cuyo objetivo es obtener estimaciones precisas de lo que esperamos ocurra en la población. El primer paso se concreta en la estadística descriptiva y representación gráfica de variables, que se mostrará en este capítulo; el segundo paso, que incluye la inferencia estadística, se abordará en un próximo capítulo de este libro.

Como hilo conductor de estos capítulos recurriremos a los datos de un estudio realizado en niños y adolescentes de un centro de salud, en el que se analizaba la asociación entre el uso de teléfonos móviles, la calidad del sueño y la obesidad.

1.- Estadística descriptiva e inferencia estadística.

La estadística desarrolla los procedimientos para la recogida, depuración y análisis de datos, que tratan de ayudar a conocer la realidad y a tomar las mejores decisiones, en presencia de incertidumbre.

Dentro de la estadística podemos diferenciar dos tipos: la estadística descriptiva y la inferencial. La estadística descriptiva se limita a describir y analizar un conjunto de datos, mientras que la estadística inferencial (o inductiva) trata de sacar conclusiones sobre una "población" a partir del análisis de los datos extraídos de un subconjunto de la misma ("muestra"). La inferencia estadística es el objetivo principal de la estadística, ya que nos permite cuantificar nuestra incertidumbre, estimando la probabilidad de error en cualquier decisión. Dentro de la inferencia estadística, podemos diferenciar a su vez dos tipos de estrategias: la estimación de intervalos de confianza de nuestras estimaciones y el contraste de hipótesis, en el que confrontamos dos o más alternativas, cuantificando la probabilidad de que las diferencias entre ellas se deban al azar.

Veamos un ejemplo para ilustrar las distintas funciones de la estadística, empleando los datos de nuestro estudio sobre uso de móviles. En un grupo de 214 niños o adolescentes se recogió el número de horas diarias de uso de móvil clasificando como abuso de móvil cuando se usaban más de dos horas. Los niños con abuso de móvil durante la semana tenían con más frecuencia obesidad (34/113) que los que no (18/101). La estadística descriptiva nos permite estimar el porcentaje de obesidad con y sin abuso de móvil, de 30% y 17,8%, respectivamente, con una diferencia de 12,3% entre grupos.

La inferencia estadística, mediante estimación de intervalos, nos facilita avanzar hasta estimar que con un 95% de confianza (5% de error) la diferencia observada del 12,3% se situaría en la población en un intervalo entre el 1% y el 23,5%, como vemos, siempre mayor en el grupo con abuso de móviles. Por último, el contraste de hipótesis

nos lleva a calcular (test de Ji-cuadrado) que la probabilidad de que las diferencias encontradas sea debida al azar es 0,037 (3,7%); como este error es muy bajo (por convención, menor de 0,05 o 5%) asumimos que el riesgo de obesidad se asocia al abuso de móviles.

2.- Estadística Descriptiva.

El primer paso del análisis estadístico es el cálculo de medidas descriptivas de la muestra de estudio. Podemos diferenciar varios grupos de medidas: de masa, de tendencia (o centralización) y de dispersión.

2.1.- Medidas de Masa:

Las medidas de masa describen la magnitud de los datos. Las más utilizadas son el tamaño muestral (n), el sumatorio y las frecuencias absoluta y relativa.

- Tamaño muestral: el tamaño muestral es el recuento del número de casos de la muestra.
- Sumatorio: el sumatorio de una variable (X_i), representado habitualmente con el símbolo Σ es la suma aritmética de los valores de una variable en todos los casos.
- Frecuencia absoluta: la frecuencia absoluta es el recuento del número de ocurrencias de cada valor de una variable.
- Frecuencia relativa: la frecuencia relativa es la proporción de casos con un valor de una variable respecto del total de casos; se puede expresar como proporción o porcentaje.

En la tabla 1 se pueden ver las frecuencias absolutas, relativas y acumuladas de la variable número de horas diarias de uso de móvil de 214 escolares, tal y como lo ofrecen las mayoría de los paquetes estadísticos.

2.2.- Medidas de posición o tendencia o centralización

Las medidas de posición, también conocidas como de tendencia o de centralización, describen la posición o magnitud alrededor de la

cual se sitúan la mayoría de los datos o bien, los valores más frecuentes. Las principales medidas de tendencia son la media, la moda y la mediana. Cada una de ellas describe una característica de los datos que estamos analizando.

Media muestral: la media muestral es la media aritmética del conjunto de valores de una variable. Se calcula dividiendo el sumatorio de los valores por el tamaño muestral.

$$\overline{X} = \frac{X_1 + X_2 + ... + X_N}{n} = \frac{\sum_{i=1}^{n} Xi}{n} = \frac{\sum Xi}{n}$$

Donde X1 a Xn son los valores cada uno de los datos, Σ el sumatorio y n el tamaño muestral. Empleando los datos de la tabla 1, podemos estimar la media dividiendo el sumatorio (719,3) por el tamaño muestral (214), con lo que obtenemos una media de 3,36 horas/día.

Moda muestral: la moda es el valor que más se repite (puede no existir y si existe puede no ser única). En la tabla 1 vemos que la moda es 2 horas, que tienen 39 casos, el 18,2% del total.

Mediana: si ordenamos los valores de una variable de menor a mayor (Xi), la mediana es el valor que está en el medio o, si la muestra es par, la media de los valores que están en el medio. Corresponde al valor:

$$\widetilde{X} = \left\{ \begin{array}{l} X_{\frac{n+1}{2}} \text{ si } \underline{n} \text{ impar} \\ \frac{X_{\frac{n}{2}} + X_{\frac{n+1}{2}}}{2} \text{ si } \underline{n} \text{ par} \end{array} \right\} \{X_i\} ordenados$$

Donde $X_{\frac{n}{2}}$ y $X_{\frac{n+1}{2}}$ son los valores de la variable X en posición central. La tabla 1 incluye una columna con porcentajes acumulados (porcentaje de casos con cada valor o menor) que facilita la identificación de la mediana, aquella fila que incluya en su porcentaje acumulado el 50%.

La medida más popular y empleada es la media, sin embargo, cuando los valores de una muestra no siguen una distribución normal, no es una buena medida de tendencia. En estas circunstancias recomendamos utilizar la mediana. Si la media y la mediana son muy diferentes, es poco probable que el valor medio describa la tendencia de los datos (probablemente no tengan una distribución de Gauss o normal), por lo que tendremos que dar la mediana o ambas.

2.3.- Medidas de dispersión

Las medidas de tendencia no permiten describir todos los datos de una muestra, porque no informan de la dispersión de cada valor respecto del valor central. Para ello, disponemos de las medidas de dispersión, fundamentalmente, el rango, la varianza, la desviación típica, el coeficiente de variación y el rango intercuartílico.

Rango: si ordenamos los valores de menor a mayor, es la diferencia entre los valores extremos (mínimo y máximo). Puede describirse con los valores extremos o con la diferencia entre ellos:

{Xi} ordenados » Xn – X1 (Máximo - Mínimo)

Para los valores de la tabla 1 los valores menor y mayor son 0 y 16 y su rango 16.

Varianza: la varianza es la media de las diferencias entre cada valor y la media al cuadrado. Se elevan al cuadrado para evitar que las diferencias negativas se anulen con las positivas. Se representa con s^2.

$$s^2 = \frac{\sum (X_i - \overline{X})^2}{n}$$

Donde s^2 es la varianza $\sum (X_i - \overline{X})^2$, es el sumatorio de las diferencias de cada valor menos la media al cuadrado y n el tamaño muestral. El cálculo lo facilitan los paquetes estadísticos e incluso las hojas de cálculo. Si lo hiciéramos manualmente tendríamos que restar de cada valor (por ejemplo 16) la media (3,36), con lo que obtenemos

la diferencia (16-3,36=12,64), que se eleva al cuadrado (159,76), sumar los cuadrados obtenidos para cada valor (si un valor se repite se multiplica por su frecuencia absoluta) y dividir por el tamaño muestral. Para los datos de la tabla 1 la varianza es 7,27.

Cuasivarianza: la cuasivarianza es una fórmula de estimación corregida de la dispersión de los datos. Aunque la varianza describe fielmente la dispersión de los datos de la muestra, si ésta tiene pequeño tamaño muestral, infraestima la dispersión de los datos en la población; por ello la fórmula se corrige disminuyendo su denominador. La varianza que se emplea en inferencia estadística es la cuasivarianza, también conocida como varianza muestral o estimada o simplemente varianza.

$$s^2 = \frac{\sum \left(X_i - \overline{X}\right)^2}{n-1}$$

Desviación típica muestral o estimada: Como en el cálculo de la varianza las distancias entre cada valor y la media se elevan al cuadrado la magnitud de la dispersión pierde sentido. Por ello, recurrimos a redimensionar la dispersión haciendo la raíz cuadrada de la varianza. De ahí resulta la desviación típica, también llamada desviación estándar, representada por "s".

$$s = +\sqrt{s^2} = \sqrt{\frac{\sum \left(X_i - \overline{X}\right)^2}{n-1}}$$

Para los datos de la tabla 1 la desviación típica es 2,69. Nos indica que cada caso de la muestra presenta un valor situado a una distancia promedio respecto a la media de 2,69 horas.

Al igual que con la varianza, existe una fórmula no corregida o desviación típica poblacional, que es:

$$s = +\sqrt{s^2} = \sqrt{\frac{\sum \left(X_i - \overline{X}\right)^2}{n}}$$

Coeficiente de variación: el coeficiente de variación estima la dispersión de los datos como medida ajustada. Al dividir la desviación típica por la media, nos indica la proporción de dispersión con respecto a la media. Generalmente se expresa como porcentaje. Resulta útil para comparar el grado de dispersión de variables con distintas unidades de medida o rango.

$$CV = \frac{s}{\overline{X}} \rightarrow CV = \frac{s}{\overline{X}} x100$$

Donde CV es el coeficiente de variación, s la desviación típica y $\overline{X}$ a media. Para los datos de la tabla 1, el coeficiente de variación es 80% (2,69/3,36x100)

Percentiles 25-75: si ordenamos los valores de una variable (X_i) de menor a mayor, el valor que deja a su izquierda el 25 % de los casos es el Percentil 25 y el que deja por arriba a un 25% de los casos el Percentil 75. El Rango intercuartílico es el intervalo entre ambos percentiles. En muestras con distribución no normal es la mejor alternativa a la desviación típica, como medida de dispersión. Aunque el rango intercuartílico lo calcula automáticamente cualquier paquete estadístico, las frecuencias de la tabla 1 facilitan su identificación usando los porcentajes acumulados. Serán aquellos valores que incluyan en su porcentaje acumulado el 25% y el 75%, en este caso 1,5 y 5; puede informarse con estos valores o con su diferencia 3,5.

3.- Métodos de Representación Gráfica.

La representación gráfica de los resultados científicos es una herramienta de gran utilidad pero si no se realiza correctamente puede inducirnos a error. Por ello, se recomienda no utilizar recursos gráficos que distraigan la atención de los valores representados y no manipular las escalas para magnificar diferencias no existentes.

Según el tipo de variable a representar, principalmente su escala de medida, disponemos de distintos diagramas:

Variables cuantitativas continuas (si el número de posibles valores es infinito, por ejemplo, el peso o la talla): el histograma, el diagrama de tallo y hojas, el diagrama de cajas y el diagrama de dispersión. Para representar la relación entre dos variables cuantitativas continuas tenemos el diagrama de puntos.

- Variables cuantitativas discretas (si el número posible de valores es finito, por ejemplo, número de hijos en una familia): el diagrama de líneas o barras.
- Variables cuantitativas o cualitativas ordinales (valores ordenados sin unidad constante, por ejemplo la escala de Apgar o el grado de reflujo vesicoureteral): diagrama de barras.
- Variables cualitativas nominales (por ejemplo el sexo o la raza): diagrama de sectores y diagrama de barras.

Describiremos a continuación las representaciones gráficas más comunes.

- Histograma.
- Se construye representando en el eje de abscisas (x) los valores de la variable continua, que se agrupan en intervalos, sobre ellos se dibujan tantos rectángulos como intervalos haya, cuya área debe ser proporcional a la frecuencia de casos en ese intervalo. En el eje de ordenadas (y) se indican las frecuencias absolutas o relativas de cada intervalo. La amplitud de los intervalos no tiene que ser constante pero la mayoría de las veces lo es, en este caso las alturas de los rectángulos serán proporcionales a su frecuencia. En la figura 1 se muestra el histograma de la variable tiempo de uso diario de móviles en horas (figura 1). Los valores se han agrupado en 8 intervalos; el que tiene mayor frecuencia es el intervalo de 0 a 2 horas, con 100 casos.
- Diagrama de cajas (Box-plots).

- El diagrama de cajas describe medidas de tendencia y dispersión de variables continuas de forma global o por grupos. Representa la distribución de los datos mediante una caja cuyo límite superior es el percentil 75 o tercer cuartil (Q3), el inferior es el percentil 25 o primer cuartil (Q1) y el centro es la mediana (percentil 50). Desde los extremos de la caja se prolongan unos "bigotes" que son los limites superior e inferior de la distribución, equivalentes a 1,5 veces el rango intecuartílico (Q1-Q3). Los valores que superan esos límites se denominan valores alejados (outliers). En la figura 2 representamos el diagrama de cajas de la variable horas diarias de uso de móvil en la semana, separada por sexos. El gráfico informa de la tendencia de los datos y de la asimetría de la distribución, muestra los valores máximo y mínimo e identifica de los valores atípicos, que aparecen representados como puntos con etiquetas de número de caso; habrá que comprobar si esos valores atípicos son reales o errores. Podemos ver que hay diferencias entre sexos, con más horas de uso de móvil en las niñas.

- Diagramas de líneas o de barras.

- Son representaciones de las frecuencias de valores de variables discretas u ordinales. A diferencia del histograma no necesitan ser agrupados para su representación, aunque, si el número de categorías es alto, podrían agruparse. En el eje de abscisas (x) se sitúan las etiquetas de cada valor; si la variable es ordinal, manteniendo el orden natural. Sobre esas etiquetas se dibujan líneas o barras cuya altura es proporcional a la frecuencia, según la escala indicada en el eje de ordenadas (y). En la figura 3 se presenta un diagrama de barras de la variable nivel de estudios maternos. Observamos cómo la categoría de mayor frecuencia es la de estudios primarios completos. No deberíamos caer en el error de establecer límites en el eje de

ordenadas para minimizar o magnificar las diferencias entre categorías. Para variables ordinales es la forma más adecuada de representación, ya que un diagrama de sectores impide ver la jerarquía de valores de las variables.

- Diagrama de Sectores (Quesitos):

- Los diagramas de sectores describen frecuencias relativas de variables cualitativas. En un círculo se distribuyen las categorías por sectores, proporcionales en tamaño a la frecuencia de cada una de ellas. La categoría con mayor frecuencia tendrá mayor área del círculo, correspondiendo un número de grados de la circunferencia proporcional a la frecuencia relativa (proporción x 360 grados). En la figura 4 se presenta la frecuencia de casos por sexos, con un predominio de participantes de sexo femenino.

- Diagrama de puntos:

- Los diagramas de puntos (Scatter - Plots) representan los valores de dos variables, habitualmente cuantitativas continuas, para cada caso de la muestra. En el área de un plano delimitado por dos rectas perpendiculares se dibujan puntos para cada observación, situados en una posición relativa al valor de una variable en el eje de abscisas (x) y de la otra variable en el eje de ordenadas (y). Junto a las líneas de coordenadas se sitúan las etiquetas de las variables y su escala de valores. La distribución de la nube puntos muestra la relación entre las variables. Cuando la nube tiene una tendencia creciente, los valores altos de una variable tienden a tener valores altos de la otra y existe una correlación positiva. Cuando la nube tiene una tendencia decreciente, los valores altos de una variable tienden a tener valores bajos de la otra y existe una correlación negativa. En un próximo capítulo presentaremos los estimadores que describen esta correlación. En la figura 5

se muestra la correlación entre los valores de horas diarias de uso de móvil y la demora en acostarse (en minutos respecto las 23:00 horas). Como vemos la nube sugiere una correlación positiva.

Se pueden hacer una serie de recomendaciones generales para la elaboración de los gráficos:

- Por convenio, la frecuencia de las variables debe ser proporcional a las áreas que se muestran. Sólo serán proporcionales a la altura, si todas las categorías tienen la misma anchura.
- Para fines comparativos, es mejor usar frecuencias relativas.
- Cuando los datos no son categóricos el número de clases o grupos no debe de ser ni muy grande ni muy pequeño. Generalmente se recomienda un número entre 5 y 20 categorías de igual tamaño. Otros autores recomiendan , siendo n el tamaño muestral.
- Los límites de las categorías de histogramas no deben coincidir con valores posibles de los datos (asignar como límite una cifra decimal más a la expresada en los valores, por ejemplo, para valores con 2 decimales una categoría de valores entre 0 y 1,999 y la siguiente entre 2 y 3,999).

Bibliografía:

ALTMAN DG, BLAND JM. Variables and parameters. BMJ. 1999; 318:1667.

ALTMAN DG. Practical statistics for medical research. London; Chapman & Hall,1991.

ARGIMÓN PALLÁS JM, JIMENEZ VILLA J. Métodos de investigación clínica y epidemiológica. Barcelona: Elsevier, 2006.

MILTON JS. Estadística para biología y ciencias de la Salud. México, McGraw-Hill, 2001.

NORMAN GR, Streiner DL. Bioestadística. México: Mosby/Doyma Libros, 1996.

OCHOA SANGRADOR C. Diseño y análisis en investigación. Madrid: International Marketing & Comunication, S.A., 2019.

OCHOA-BREZMES J, RUIZ-HERNÁNDEZ A, BLANCO-OCAMPO D, GARCÍA-LARA GM, GARACH-GÓMEZ A. Mobile phone use, sleep disorders and obesity in a social exclusion zone. An Pediatr (Engl Ed). 2023;98(5):344-352.

ORTEGA PÁEZ E, OCHOA SANGRADOR C, MOLINA ARIAS M. Representación gráfica de variables. Evid Pediatr. 2019;15:13.

ROSNER B. Fundamentals of Biostatistics, 7th Edition. Boston: Brooks/Cole, Cengage Learning 2011.

Tabla I.- Tablas de frecuencias de la variable número de horas/día de uso de móvil.

Horas/día	Frecuencia	Porcentaje	Porcentaje acumulado
0.0	11	5.1%	5.1%
0.1	1	0.5%	5.6%
0.2	1	0.5%	6.1%
0.3	1	0.5%	6.5%
0.5	12	5.6%	12.1%
0.7	1	0.5%	12.6%
0.8	4	1.9%	14.5%
1.0	18	8.4%	22.9%
1.5	13	6.1%	29.0%
2.0	39	18.2%	47.2%
2.5	5	2.3%	49.5%
3.0	24	11.2%	60.7%
3.5	3	1.4%	62.1%
4.0	22	10.3%	72.4%

5.0	16	7.5%	79.9%
5.5	1	0.5%	80.4%
6.0	15	7.0%	87.4%
7.0	9	4.2%	91.6%
7.5	2	0.9%	92.5%
8.0	4	1.9%	94.4%
9.0	3	1.4%	95.8%
10.0	7	3.3%	99.1%
12.0	1	0.5%	99.5%
16.0	1	0.5%	100.0%
Total			
(tamaño muestral)	214	100.0	
Sumatorio= (ΣXi)= 719,3			

Figura 1.- Histograma de frecuencias

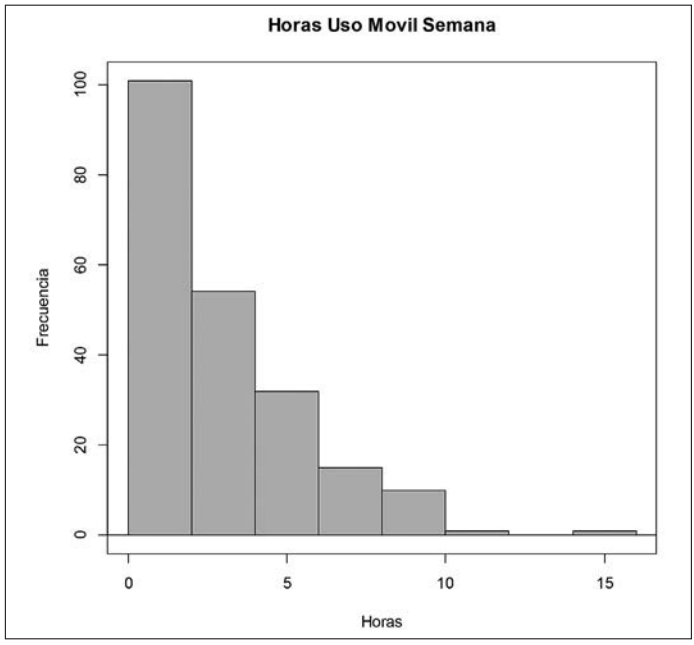
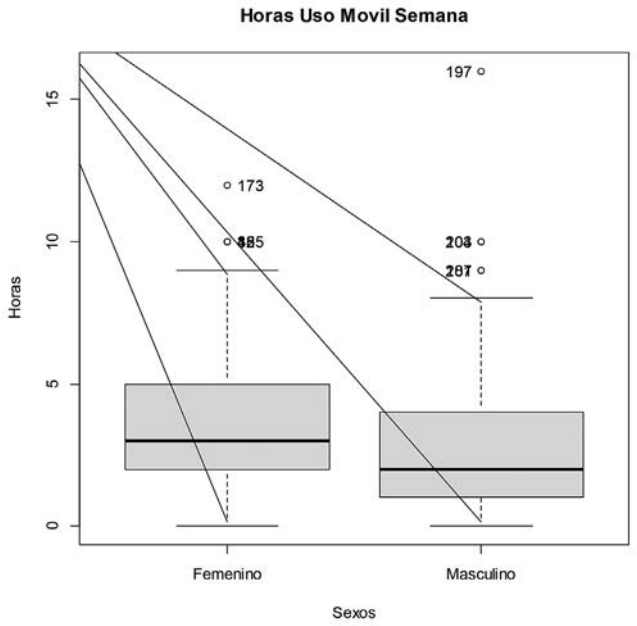
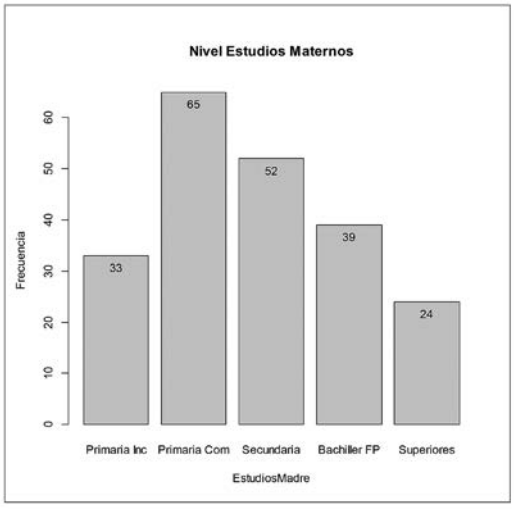


Figura 2.- Diagramas de cajas.



Mediana: línea gruesa horizontal central; Percentiles 25 y 75: límites inferior y superior de la caja sombreada; Valores extremos (descartando valores atípicos): líneas horizontales finas inferior y superior límites que comprenden los valores alejados menos de 1,5 veces el rango intercuartílico desde la mediana; Valores atípicos (señalados con el número de orden), los situados más alejados de 1,5 veces el rango intercuartílico. o >P75+1,5RIC; RIC: rango intercuartílico.

Figura 3.- Diagrama de barras.



Com: completa; FP: formación profesional; Inc: incompleta

Figura 4.- Diagrama de Sectores.

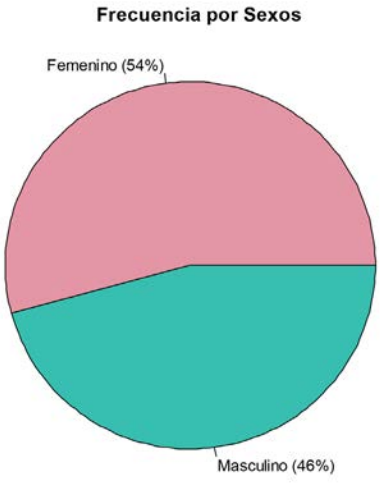
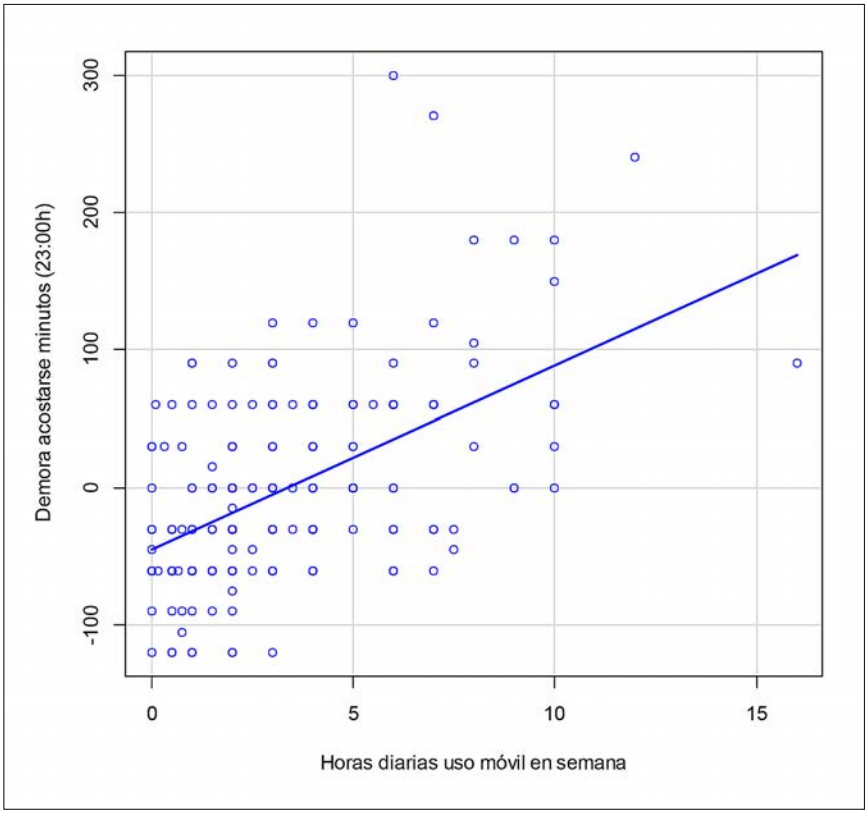


Figura 5.- Gráfico de dispersión (Scatter Plot).



Análisis avanzado de datos

Inferencia estadística.

En el apartado anterior de este libro dijimos que la estadística nos ayudaba en la toma de decisiones, ya que nos permitía cuantificar la incertidumbre, y elegir la opción más adecuada, minimizando el grado de error. En dicho capítulo se presentó la estadística descriptiva, quedando pendiente de abordar, en este, la inferencia estadística, que es la que trata de estimar a partir de los datos de una muestra, parámetros de la población de la que procede.

Dentro de la inferencia estadística diferenciamos dos tipos de estrategias:

- La estimación de intervalos de confianza, que nos informa del rango de valores entre los que se encontrará el parámetro poblacional a estimar.
- El contraste de hipótesis, con el que habitualmente confrontamos dos o más alternativas, cuantificando la probabilidad de que las diferencias entre ellas se deban al azar.

Ambas estrategias se sustentan en el cálculo de probabilidades de las variables aleatorias. Aunque este cálculo se realiza habitualmente de forma automática por los paquetes estadísticos, conviene conocer

ciertos conceptos fundamentales. Al igual que hicimos en el capítulo anterior, para ilustrar con ejemplos las explicaciones, emplearemos los datos de un estudio sobre abuso de móviles y su relación con problemas de sueño y obesidad.

1.– Probabilidad, variables aleatorias y distribuciones de probabilidad.

Uno de los objetivos habituales de la investigación es estimar la frecuencia de una determinada característica en la población, que puede ser un factor de exposición (riesgo o protección) o un efecto (enfermedad o curación). Por ejemplo, en nuestro estudio sobre abuso de móviles nos hemos planteado estimar su frecuencia en población escolar. Lo ideal sería estudiar a toda la población escolar; la frecuencia relativa obtenida sería la mejor estimación de la probabilidad de abuso de móvil en la población. Pero esta opción no es factible ni eficiente. Por ello, recurrimos a seleccionar una muestra de escolares y tratamos de estimar la frecuencia en la población a partir de la frecuencia observada en nuestra muestra.

La probabilidad de un evento se puede expresar con un número entre 0 y 1, asociado con la verosimilitud de que ocurra un suceso (0 suceso imposible; 1 suceso seguro). En nuestro estudio hemos entrevistado a 214 escolares de los que 113 referían un uso de móviles los días de diario superior a 2 horas. Podemos, pues, estimar que la probabilidad de abuso de móvil es 0,528 (113/214), expresado en porcentajes 52,8%. Si entrevistamos a un nuevo escolar, no incluido en el estudio pero de características similares, no sabemos a priori si tendrá o no abuso de móvil, pero nuestra mejor estimación de la probabilidad de que lo tenga será 0,528. Esta misma probabilidad sería aplicable a una nueva muestra de escolares que reclutáramos.

Cualquier variable medida en un estudio se caracteriza por ser fruto de mediciones repetidas de una misma característica, que se define por los posibles valores y la probabilidad de que aparezca cada

uno de ellos. Cuando los distintos valores de una variable siguen una distribución de probabilidad la denominamos variable aleatoria. Para la variable abuso de móvil los valores posibles son sí y no, y sus probabilidades respectivas son 0,528 y 0,472. La distribución de probabilidades aplicable a esta variable nominal dicotómica la conocemos como distribución binomial. Existen otras distribuciones de probabilidad, la más conocida de ellas es la distribución normal, que es la que siguen ciertas variables cuantitativas continuas, como puede ser la talla estandarizada de nuestros escolares (figura 1).

La distribución normal viene caracterizada por su simetría, por su valor central (μ, valor esperado, media o esperanza matemática) y por su dispersión, que es proporcional a la varianza (2, varianza). Nos basta con conocer la media y la varianza para estimar la probabilidad de cualquier rango de valores. Así, sabemos que a cada lado de la media se sitúan el 50% de los valores, que entre la media menos y más una unidad de desviación típica (raíz cuadrada de la varianza) se encuentran el 68% de los valores y que entre la media menos y más 1,96 veces la desviación típica están el 95% de los valores. También sabemos que, de forma asimétrica, a un solo lado de la media más (o menos) 1,65 veces la desviación típica se encuentran el 5% de los valores (figura 2).

Con estas referencias y otras disponibles en tablas detalladas podemos saber la probabilidad de cualquier rango de valores y de forma inversa el rango de valores que corresponde a una determinada probabilidad. Así podremos saber la probabilidad de cualquier resultado, si sabemos la distribución de probabilidad de referencia. Las tallas de nuestros escolares representados en la figura 1 se han estandarizado (restadas las medias y divididas por las desviaciones estándar por edad y sexo, de las tablas de crecimiento nacionales). Así vemos que entre -2 y +2 se encuentran la mayoría de los valores (alrededor del 95%), cualquier valor más alejado de estos límites será muy poco probable.

Además de las distribuciones de probabilidad ya mencionadas, binomial y normal, hay muchas otras, como la distribución de Poisson

(eventos raros que tienen lugar a lo largo de un período de tiempo o espacio), Ji cuadrado (que siguen los valores observados y esperados de una tabla de contingencia), t de Student, F de Snedecor (que siguen los cocientes de varianzas), etc. Cada una de ellas será de aplicación para diferentes test estadísticos.

2.- Estimación por intervalos. El error estándar.

Como dijimos al comienzo del capítulo la primera estrategia de la inferencia estadística era la estimación por intervalos de confianza. Para ello contamos con los parámetros estimados en nuestra muestra, habitualmente una frecuencia relativa, para variables nominales dicotómicas (ej. proporción de abuso de móvil), o una media, para variables continuas (ej. media de talla estandarizada). Nuestro objetivo es estimar las características correspondientes en la población, que se representan por las letras griegas: π (proporción) y μ (media). Como nuestras estimaciones han sido obtenidas de muestras, por prudencia, solo podemos decir que los verdaderos parámetros a estimar, por ejemplo, la proporción (π) o la media (μ) poblacionales, tendrán valores cercanos a los obtenidos en nuestra muestra.

Pero, ¿cuánto de cercanos? ¿cómo calculamos el intervalo de error alrededor de nuestras estimaciones muestrales? La aproximación más intuitiva es asumir que nuestra muestra sólo es una de las teóricas muestras que podríamos haber estudiado, por lo que nuestras estimaciones de proporción o media, son sólo una de las estimaciones posibles. Sabemos que si estudiáramos sucesivas muestras del mismo tamaño el conjunto de las estimaciones obtenidas (ej. proporción o media) en cada una ellas siguen una distribución normal, para tamaños muestrales $\geq$30 (teorema central del límite). Esta distribución tiene una desviación estándar inversamente proporcional al tamaño muestral, a mayor tamaño muestral menor dispersión. La desviación estándar de las estimaciones muestrales se conoce como error estándar o error típico, que será el que utilicemos para la estimación de intervalos de confianza y los contrastes de hipótesis.

La distribución de estimaciones de proporciones se aproxima aceptablemente a una normal cuando el producto tamaño muestral por la proporción y por su complementario (n·p·[1-p]) es mayor de 5. Esta distribución normal tiene un valor esperado, que es la propia proporción, y una desviación estándar, que ahora denominamos error estándar, que se calcula con la fórmula:

$$\text{Error Estándar}_{proporción} = \sqrt{\frac{\pi \cdot (1-\pi)}{n}} \approx \sqrt{\frac{p \cdot (1-p)}{n}}$$

En la que π es la proporción poblacional, p la proporción muestral y n el tamaño muestral

Vemos que para el cálculo del error estándar necesitaríamos saber la proporción poblacional, desconocida, por lo que usaremos la proporción de nuestra muestra, que será la estimación puntual de nuestro parámetro. Con esta estimación y el tamaño muestral podemos calcular el error estándar. Para nuestra estimación de frecuencia de abuso de móvil el cálculo sería:

$$Errorestándar_{proporción} = \sqrt{\frac{0{,}528 * 0{,}472}{214}} = 0{,}034$$

Ahora, si utilizamos las propiedades de la distribución normal, podemos cuantificar entre qué rango de valores puede encontrarse la proporción poblacional. Recordemos que en un rango entre 1,96 veces por debajo y por arriba de la media se encontraban el 95% de los valores. Usemos esta propiedad de la distribución normal para estimar el rango de valores entre los que tengo una confianza del 95% de que se encuentre la proporción poblacional. Este cálculo corresponde a:

$$p \pm Z_{1-\alpha/2} \cdot \text{error estándar} \Rightarrow \text{para } 95\% \text{ confianza} \Rightarrow 0{,}528 \pm 1{,}96 \cdot 0{,}030$$

Para nuestra proporción y una confianza del 95% el intervalo de la proporción poblacional (p) estará entre 0,462 y 0,594 (46,2%

a 59,4%). Habitualmente este intervalo se conoce como Intervalo de Confianza del 95 % (IC95%). Se puede interpretar diciendo que tenemos un 95% de confianza de que la proporción poblacional se encuentre entre 46,2% y 59,4% (con un error menor del 5%). Aunque esta es la interpretación más sencilla, la interpretación real corresponde a que si hiciéramos 100 estudios con un tamaño muestral similar al nuestro, el verdadero parámetro poblacional estaría incluido en 95 de los 100 intervalos de confianza estimados.

Al igual que hemos presentado la fórmula del error estándar para la estimación de proporciones, existen otras fórmulas de error estándar de otros parámetros (figura 3). El procedimiento siempre es el mismo: a) calculamos la medida descriptiva o la medida de frecuencia, riesgo o impacto de nuestra muestra; b) estimamos a partir de dichos datos el error estándar; y c) usamos dicho error estándar y el valor Z correspondiente para estimar el intervalo de confianza (para un IC95%: 1,96).

Aunque podríamos hacer un cálculo manual de intervalos de confianza usando las fórmulas de los errores estándar no lo recomendamos, ya que distintos paquetes estadísticos y calculadoras epidemiológicas realizan dichos cálculos. Podemos mencionar como calculadoras epidemiológicas de acceso libre, Epidat, en versión instalable de escritorio (https://www.sergas.es/Saude-publica/EPIDAT-4-2?idioma=es), y Calcupedev en versión online (https://www.aepap.org/calculadora-estudios-pbe/#/).

Empleando Calcupedev, en concreto la hoja para estudios transversales (https://www.aepap.org/calculadora-estudios-pbe/#/estudios-transversales), podemos estimar la diferencia de proporciones de obesidad entre los escolares con y sin abuso de móvil durante la semana. En nuestro estudio encontramos que 34 de 113 niños (30%) con abuso de móvil tenían obesidad, frente a 18 de 101 (17,8%) sin abuso de móvil. Introduciendo en la tabla tetracórica los valores con y sin obesidad en cada grupo (34 y 79 frente 18 y 83) obtenemos que la

diferencia (ver medidas de impacto) es 0,123 (12,3%), con un intervalo de confianza del 95% entre 0,01 y 0,235 (1% y 23,5%). Como vemos, el intervalo solo incluye porcentajes positivos, que implican mayor riesgo en los que tienen abuso de móvil, por lo que podemos intuir, con un 95% de confianza, que hay asociación entre abuso de móvil y obesidad. No obstante, este razonamiento ya corresponde al contraste de hipótesis, que veremos a continuación.

3.- Contraste de Hipótesis.

Recordamos que el contraste de hipótesis permite comparar dos o más muestras o estimaciones muestrales, cuantificando la probabilidad de que las diferencias entre ellas se deban al azar.

En el ejemplo que veíamos antes, queríamos saber si el abuso de móvil se asociaba a mayor riesgo de obesidad. En el contraste de hipótesis se plantean dos alternativas:

- Hipótesis nula: no hay asociación entre abuso de móvil y obesidad. La diferencia de proporciones no es distinta de 0.
- Hipótesis alternativa; dos opciones: a) sí hay asociación entre abuso de móvil y obesidad, positiva o negativa (contraste bilateral), o b) el abuso de móvil se asocia a mayor riesgo de obesidad (contraste unilateral). La diferencia de proporciones es distinta (opción bilateral) o mayor (opción unilateral) que 0. La elección de un contraste bilateral o unilateral depende del investigador, el bilateral es el más conservador, pero ambos pueden ser válidos.

Anteriormente habíamos estimado la diferencia de proporciones de obesidad, observando que aquellos con abuso de móvil tenían un 12,3% mayor riesgo de obesidad, con un IC95% entre 1% y 23,2%. Como ese intervalo no incluye el valor nulo, que para una diferencia es "0", podíamos ya asumir la existencia de asociación. Sin embargo, para resolver el contraste de hipótesis debemos cuantificar la probabilidad exacta de que la diferencia encontrada sea mayor que "0" por azar.

Contamos con varias pruebas con las que calcular esta probabilidad. Una de las opciones es la aproximación a la distribución normal de la diferencia de proporciones, cuyo error estándar (Figura 3) era:

$$EE_{\substack{Diferencia\\proporciones}} = \sqrt{\frac{p_1 \cdot (1-p_1)}{n_1} + \frac{p_2 \cdot (1-p_2)}{n_2}}$$

Ahora podemos calcular cuál es la probabilidad de obtener una diferencia de 0,123, bajo la asunción de la hipótesis nula de que no haya diferencias, o lo que es lo mismo, que la diferencia sea 0 (elegimos una hipótesis bilateral de no existencia de diferencias). La probabilidad de la diferencia se estima calculando el valor Z estandarizado correspondiente, con la fórmula:

$$Z = \frac{p_1 - p_2}{\sqrt{\frac{p_1 \cdot (1-p_1)}{n_1} + \frac{p_2 \cdot (1-p_2)}{n_2}}} = \frac{0{,}300 - 0{,}178}{\sqrt{\frac{0{,}300 \cdot (1-0{,}300)}{113} + \frac{0{,}178 \cdot (1-0{,}178)}{101}}} = 2{,}089$$

Ahora basta conocer la probabilidad del valor Z (2,089). Si comparamos este valor con los de referencia anteriormente mencionados, vemos que es superior a 1,96 (límite del 95% bilateral); por ello, suponemos que la probabilidad de ese valor será menor del 5% (0,05). Pero ¿cuál es exactamente esa probabilidad? Aunque contamos con tablas de referencia, lo habitual es que recurramos a paquetes estadísticos o calculadoras. Si introducimos los datos en Epidat, en el menú Módulos-Inferencia sobre parámetros-Dos poblaciones-Proporciones independientes, el programa nos da el valor Z, que coincide con el calculado por nosotros (2,089), y su probabilidad, que es 0,037 para un contraste bilateral. Esta probabilidad (p) es muy baja, lo suficientemente baja como para que si asumimos que hay asociación entre abuso de móvil y obesidad la probabilidad de equivocarnos sea 0,037.

Como esta probabilidad es menor de 0,05 (5%), por convención, podemos rechazar la hipótesis nula (no hay asociación) y aceptar la alternativa (si hay asociación). Rechazando la hipótesis nula y acep-

tando la alternativa asumimos un error de 0,037. A este error lo denominamos error tipo I, o error de falso positivo (encontrar diferencias en la muestra cuando no las hay en la población), y a su probabilidad la llamamos "alfa". Es importante advertir que, aunque el error sea muy pequeño, siempre existe cierto riesgo de error.

En nuestro ejemplo anterior hemos rechazado la hipótesis nula, pero ¿qué hubiera pasado si la muestra de estudio hubiera sido de menor tamaño? Imaginemos un estudio que incluyera sólo 90 escolares con abuso de móvil y 89 sin él; de ellos tenían obesidad 27 del primer grupo y 16 del segundo, lo que suponen proporciones y diferencias similares a los observados con la muestra previa. Al disminuir el tamaño muestral, el error estándar aumentaría (al estar en la denominador de la fórmula), el valor Z disminuiría y su probabilidad aumentaría. Si introducimos los nuevos datos en la calculadora Epidat, Z resulta 1,882 (menor de 1,96) y su probabilidad 0,06 (mayor de 0,05). Por lo tanto, con esta nueva muestra no podríamos rechazar la hipótesis nula ni aceptar la alternativa, ya que el error tipo I (o de falso positivo) en el que incurriríamos sería mayor de 0,05 (5%).

¿Qué ha pasado? Que el nuevo estudio ha perdido potencia respecto al anterior, aumentando el riesgo de error tipo II (riesgo beta) o de falso negativo (probabilidad de no encontrar diferencias en la muestra cuando sí las hay en la población). Podría existir asociación entre abuso de móvil y obesidad, pero no hemos podido encontrarla.

Para que un resultado "negativo" (no hay diferencias estadísticamente significativas) sea aceptable el estudio debe tener un tamaño muestral lo suficientemente grande como para detectar diferencias que sean clínicamente importante. Esto se concreta en que la probabilidad de no encontrar diferencias sea baja, esto es, que el error tipo II sea bajo, lo que por convención se considera si es menor de 0,20 (20%). Al complementario del riesgo beta lo llamamos Potencia (1-beta). Por ello, un resultado "negativo" solo es aceptable si la potencia es mayor del 80%. Para calcular el error tipo II (riesgo beta) o la potencia (1-beta), reco-

mendamos usar una calculadora epidemiológica, como Epidat (menú Módulos-Muestreo-Cálculo de tamaño de muestra-Contraste de hipótesis-Comparación de proporciones-Proporciones independientes). Con una muestra de 179 sujetos, la potencia para detectar diferencias de al menos 12% entre grupos es del 3,7%, como vemos, muy inferior al 80% necesario; por lo tanto necesitaríamos mayor tamaño muestral para poder asumir la ausencia de diferencias. Es importante destacar que en el cálculo de la potencia del estudio debemos introducir diferencias que consideremos clínicamente importantes, que no tienen por qué coincidir con las observadas en nuestro estudio.

En la tabla I se resumen todas las situaciones posibles del contraste de hipótesis. Debemos tener en cuenta que sea cual sea la decisión de nuestro contraste, siempre existe un cierto riesgo de error, ya que la población es inaccesible.

4.- Elección de la prueba de contraste de Hipótesis.

Anteriormente hemos realizado contraste de hipótesis empleando la aproximación a la normal de la diferencia de proporciones, pero existen muchas otras pruebas, entre las que tendremos que elegir la más apropiada para cada contraste.

En la elección del test estadístico tendremos que considerar: el número de variables implicadas, cuáles son las variables dependiente e independiente, qué escalas de medidas tienen las variables (nominal, ordinal, continua normal, continua no normal), la existencia o no de dependencia entre grupos comparados, el tipo de contraste (uni o bilateral) y los umbrales de errores tipo I y II (habitualmente 0,05 y 0,20).

En la tabla 2 se presenta un esquema para la elección de la prueba más apropiada. Las columnas diferencian en función de la escala de la variable dependiente, mientras que las filas diferencian según la escala de la variable independiente, que en sus primeras filas corresponde a la comparación de dos grupos (variable independiente nominal dicotómica), independientes o relacionados, o más de dos grupos

(variable independiente nominal politómica). Sin embargo, la última fila corresponde a cuando la variable independiente es continua, por lo tanto, no clasifica grupos. El esquema puede ser interpretado de forma invertida, buscando en las columnas la variable independiente y en las filas la dependiente, sin que ello afecte a la adecuación de la prueba elegida. Además de estas pruebas, existen otras que no aparecen en el esquema, que mencionaremos más adelante.

Queda fuera del objeto de este capítulo repasar pormenorizadamente cada una de las pruebas de contraste de hipótesis, por lo que nos centraremos en las condiciones de aplicación e interpretación de las más habituales: la prueba de Ji cuadrado, la t de Student, el análisis de la varianza (ANOVA) y los coeficientes de correlación; incluiremos también alguna prueba no paramétrica. Para ello, vamos a recurrir a los resultados del estudio anteriormente mencionado. En la tabla 3 se presentan los análisis descriptivos y contrastes de una selección de variables, que incluyen el nivel de significación (p) y la prueba elegida en cada ocasión.

4.1.- Variables nominal dicotómica frente a nominal dicotómica: Comparación de proporciones entre dos grupos.

En la mitad superior de la tabla se muestran algunas variables que podrían estar asociadas al abuso de móvil durante la semana. Para elegir la prueba a aplicar debemos buscar en nuestro esquema (tabla 2) las variables implicadas: la variable abuso de móvil, que podemos considerar variable dependiente, es nominal, por lo que nos situaremos en la primera columna del interior del esquema. La primera variable a asociar con ella, el sexo, que consideraremos variable independiente, es una variable nominal dicotómica, por lo que nos situaremos en la primera fila del esquema, que corresponde a dos muestras independientes (ya que los grupos comparados, diferenciados por sexo, son independientes entre sí). Vemos que en la casilla cruzada hay varias pruebas, alguna de ellas vista previamente en este capítulo (test Z de comparación de proporciones), de las que la más usada es el test de Ji cuadrado.

Elijamos la prueba de Ji cuadrado. Esta prueba se basa en la propiedad que siguen los recuentos esperados y observados en las casillas de las tablas de contingencia y es aplicable tanto a tablas de 2x2, como la de nuestro ejemplo (sexo masculino/femenino vs abuso de móvil si/no), como a tablas con más de 2 filas o columnas (n x m), de las que más adelante se verá un ejemplo. En la tabla 3 no aparecen todos los recuentos de la tabla de contingencia, ya que se han omitido los recuentos y porcentajes de sexo femenino, aunque son fácilmente deducibles restando los masculinos del total de cada grupo. Viendo los datos podría parecer que el abuso de móvil es menos frecuente entre los varones, aunque la diferencia es discreta y la probabilidad de que se deba al azar (0,114) superior al nivel de significación (>0,05), por lo que no podemos decir que haya asociación entre abuso de móvil durante la semana y el sexo.

La siguiente variable para analizar es obesidad, que tiene las mismas características que sexo, nominal dicotómica, dos muestras independientes. Por ello elegimos de nuevo la prueba de Ji cuadrado. A la diferencia entre grupos, 12,4% más obesidad entre los que tienen abuso de móvil, le corresponde una probabilidad de 0,037, que es menor de 0,05 (no esperada por azar), por lo que concluimos que hay asociación entre abuso de móvil y obesidad, asumiendo una probabilidad de error de 0,037.

En el esquema vemos que hay una tercera prueba, el test exacto de Fisher. Esta prueba, es una alternativa a la de Ji cuadrado, cuando esta última no se puede aplicar, aunque podría emplearse en cualquier circunstancia. La prueba de Ji cuadrado requiere que no haya en ninguna casilla de la tabla 2x2 valores esperados inferiores a 5; si la tabla tiene más de dos filas o columnas podría haberlos en menos del 20% de las casillas. Los paquetes estadísticos informan de esta circunstancia y habitualmente facilitan el resultado de la prueba de Fisher, que usaremos cuando creamos conveniente.

4.2.- Variables nominal politómica frente a nominal: Comparación de porcentajes entre más de dos grupos.

La siguiente variable para analizar, que vemos en la tabla III, es el nivel de estudios maternos, agrupados en tres categorías. Hemos decidido, por conveniencia, considerarla variable nominal potitómica, aunque podría analizarse como variable ordinal; esto es así por las dudas que tenemos a la hora de cuantificar numéricamente cada categoría. En el esquema de la tabla II nos situaríamos en la primera columna para variables nominales (abuso de móvil) y en la tercera fila de variables nominales politómicas (estudios maternos). Vemos que, de nuevo, entre las opciones del esquema, aparece la prueba de Ji cuadrado. La tabla de contingencia a analizar sería una tabla de 3 filas por 2 columnas. Los porcentajes de nivel de estudios son claramente diferentes entre los escolares con y sin abuso de móvil, lo que se traduce, en la prueba de Ji cuadrado, en una probabilidad <0,001. En realidad, el paquete estadístico ofrece una probabilidad mucho menor de esa cifra, pero por acomodar la presentación se suele truncar a ese nivel de significación. Podemos concluir, pues, que hay asociación entre abuso de móvil y el nivel de estudios maternos, con una probabilidad de error muy baja.

4.3.- Variables nominal dicotómica frente a continua: Comparación de medias entre dos grupos independientes.

La siguiente variable de la tabla 3 es una variable continua, la talla de los escolares en su valor estandarizado. En la tabla nos situamos en la primera columna (variable dependiente nominal) y en la cuarta fila (variable independiente continua), encontrando como opción el test de la t de Student. Aunque esto es correcto, la elección de la prueba es más precisa si invertimos la lectura del esquema. Ya adelantamos que la interpretación es igualmente válida si intercambiamos las variables dependiente e independiente. Esta lectura ofrece más opciones, que optimizan la elección de la prueba. En este caso escogeríamos la tercera columna para la variable talla (continua) y la primera fila para la varia-

ble abuso de móvil (nominal dicotómica con muestras independientes). Vemos que en la casilla se muestra la t de Student, pero en concreto la variante para muestras independientes. Antes de proceder a la aplicación de esta prueba tenemos que comprobar si la variable talla sigue una distribución normal, porque en caso contrario, la columna a leer sería la situada a su izquierda, que es la que corresponde a las variables ordinales y a las continuas de distribución no normal.

Es fácil asumir la normalidad de una variable ya estandarizada, además en la figura 1 mostramos el histograma de esta variable, que parece adecuarse a una curva de Gauss; podríamos aplicar una prueba de comprobación de normalidad, como el test de Kolmogorov-Smirnov, que apoyaría la normalidad, pero en este caso es innecesario. Por lo tanto, escogemos un test de la t de Student para muestras independientes. En la tabla vemos que las medias de los grupos son similares y su diferencia mínima. Para calcular la probabilidad de que esa diferencia se deba al azar, debemos elegir un contraste diferente según las varianzas de cada grupo sean homogéneas o no. La probabilidad que se ha anotado en la tabla corresponde a la calculada para varianzas homogéneas. Como vemos, es muy probable (p=0,772) que la diferencia observada se deba al azar (>>0,05) y no podemos asumir asociación entre talla y abuso de móvil.

La siguiente variable de la tabla, demora al acostarse, es también una variable continua, pero no sigue una distribución normal, fundamentalmente por su asimetría (escolares con demoras elevadas fuera de rango). Tanto los métodos gráficos como los test de normalidad confirmarían nuestra sospecha. Por lo tanto en el esquema de la tabla III nos situaríamos en la segunda columna (variables ordinales o continuas no normales) y al igual que en el análisis anterior en la primera fila (nominal dicotómica para muestras independientes). La prueba a elegir es una prueba no paramétrica, el test de Mann-Whitney. En la tabla 3, en vez de medias y desviaciones típicas, hemos representado las medianas y rangos intercuartílicos, medidas mucho más adecuadas en

ausencia de normalidad. Vemos que la diferencia de demora entre grupos es importante, de 30 minutos; la probabilidad de que esa diferencia se deba al azar es muy baja, <0,001. Por lo tanto, podemos asumir que hay asociación entre abuso de móvil y demora en acostarse, con una probabilidad de error muy baja.

4.4.- Variables nominal politómica frente a continua: Comparación de medias entre más de dos grupos.

En la mitad inferior de la tabla III aparecen los descriptivos de otra variable dependiente, en este caso el índice de masa corporal (IMC). Esta variable es continua y asumiremos que es normal, por ser valores estandarizados por edad y sexo. Se presentan la media y desviación estándar por grupos de sexo y por estudios maternos. Para la variable sexo, la opción a elegir sería la misma anteriormente empleada con la talla, tercera columna para variables continuas, primera fila para variables nominales dicotómicas, lo que nos lleva a elegir la t de Student para muestras independientes.

Sin embargo, la siguiente variable, estudios maternos, nos lleva en el esquema de la tabla II a la tercera columna, para IMC, variable continua, y a la tercera fila, para estudios maternos, variable nominal politómica. La elección apropiada en este caso es el ANOVA. Además de la normalidad, el ANOVA tiene otros requisitos, como la existencia de homocedasticidad (homogeneidad de varianzas), requisito que se cumple en este caso (existen pruebas específicas para su comprobación, que omitimos). Como vemos, hay diferencias importantes en las medias, especialmente para el grupo con más nivel de estudios, que tiene menor IMC. La probabilidad de que esa diferencia se deba al azar es 0,036, poco probable y <0,05, por lo que podemos asumir que hay asociación entre nivel de estudios maternos e IMC. Como son tres grupos los comparados podría interesar saber entre qué grupos existe diferencia; disponemos de distintos análisis post-hoc que muestran los grupos con diferencias entre sí.

4.5.- Variables continua frente a continua: correlación.

En el capítulo anterior presentamos como ejemplo de diagrama de puntos una representación de los valores de tiempo de demora al acostarse y horas diarias de uso de móvil durante la semana (Figura 5 del capítulo anterior). Dijimos en su momento que la nube de puntos sugería una tendencia creciente en la que a mayor número de horas de uso de móvil había una mayor demora al acostarse. Esto es lógico, pero si queremos cuantificar dicha asociación tendremos que usar algún estadístico que la describa. En la figura se había dibujado la línea de ajuste a la nube de puntos. Cuando esta línea es plana se considera que no hay correlación, si la línea es creciente y coincide con la diagonal la correlación es máxima, lo que indica correlación positiva, lo mismo ocurre si es decreciente siguiendo la diagonal inversa, lo que indica correlación negativa. Estas tres situaciones se corresponden, cuantitativamente, con lo que conocemos como coeficientes de correlación, de 0, +1 y -1. En función de la pendiente tendremos coeficientes más o menos alejados de 0. Para saber si existe correlación debemos estimar la probabilidad de que el coeficiente sea significativamente diferente de 0 (ausencia de correlación).

Volvamos a nuestro ejemplo. Las dos variables que se quieren correlacionar son variables continuas. En el esquema de la tabla 2 nos situaríamos en la tercera columna y la cuarta fila, ambas para variables continuas. No obstante, aunque no mostraremos representaciones gráficas ni test de contraste de normalidad, es fácilmente asumible que ninguna de estas variables sigue una distribución normal (una de ellas se ha representado gráficamente en la figura 1 del capítulo anterior). Por ello, nos desplazaremos a la segunda columna, la correspondiente a variable ordinales o continuas no normales. Si ambas variables de distribuyeran normalmente hubiéramos elegido la tercera columna y la cuarta fila (correlación de Pearson), pero como no son normales escogemos la prueba de la segunda columna, cuarta fila, en la que encontramos la correlación de Spearman.

El coeficiente de correlación de Spearman es 0,50 y le corresponde una probabilidad <0,001. Así, si asumimos que existe una correlación positiva entre ambas variables, la probabilidad de error será muy baja (<<0,05).

4.6.- Otros contrastes.

En el esquema de la tabla 2 aparecen varias pruebas aplicables a la comparación de variables o grupos relacionados entre sí (principalmente en la fila 2), tanto para variables nominales, como ordinales y continuas. Cuando medimos una variable en los mismos sujetos en dos ocasiones distintas (por ejemplo, antes y después de un tratamiento) el análisis debe tener en cuenta esta relación. Si comparamos dos grupos cuyos sujetos han sido apareados por covariables de interés, también.

Además de las pruebas contenidas en la tabla 2, existen muchas otras aplicables a circunstancias particulares, omitidas por facilitar la interpretación del esquema. Así, merecen la pena destacarse los análisis de supervivencia, que permiten analizar variables tiempo-evento. Estas técnicas resultan muy útiles, no solo para analizar la supervivencia, también para otras variables como: tiempo hasta recaída, tiempo hasta abandono de lactancia materna, tiempo hasta retirada de una intervención, etc.

Con frecuencia necesitamos analizar la relación entre más de dos variables, habitualmente porque existen covariables asociadas simultáneamente con la variable dependiente en estudio. Para ello se han desarrollado diferentes técnicas de análisis multivariante, cuya explicación excede los objetivos de este capítulo. Podemos mencionar las más empleadas, que diferenciaremos en función de la escala de medida de la variable dependiente: para variables continuas la regresión lineal múltiple, para variables nominales la regresión logística múltiple, para variables tiempo-evento la regresión de Cox. Animamos al lector interesado a consultar los documentos referenciados en la bibliografía para ampliar información.

Bibliografía

ALTMAN DG, Bland JM. The normal distribution. BMJ. 1995; 310:298.

ALTMAN DG. Practical statistics for medical research. London; Chapman & Hall,1991.

ARGIMÓN PALLÁS JM, JIMENEZ VILLA J. Métodos de investigación clínica y epidemiológica. Barcelona: Elsevier, 2006.

KLEIBAUM, DG, KUPPER LL, MULLER KE. Applied Regression Analysis and other Multivariable Methods. 3rd Edition. Boston: PWS–KENT Publishing Company 1998.

MILTON JS. Estadística para biología y ciencias de la Salud. México, McGraw-Hill, 2001.

NORMAN GR, STREINER DL. Bioestadística. México: Mosby/ Doyma Libros, 1996.

OCHOA SANGRADOR C. Diseño y análisis en investigación. Madrid: International Marketing & Comunication, S.A., 2019.

OCHOA-BREZMES J, RUIZ-HERNÁNDEZ A, BLANCO-OCAMPO D, GARCÍA-LARA GM, GARACH-GÓMEZ A. Mobile phone use, sleep disorders and obesity in a social exclusion zone. An Pediatr (Engl Ed). 2023; 98(5):344-352.

RIEGELMAN RK, HIRSH RP. Cómo estudiar un estudio y probar una prueba: lectura crítica de la literatura médica. 2a. ed. Washington, D.C. Organización Panamericana de Salud. 1992. (Publicación Científica nº 531)

Rosner B. Fundamentals of Biostatistics, 7th Edition. Boston: Brooks/Cole, Cengage Learning 2011.

Tabla I.- Alternativas del contraste de hipótesis.

Decisión	Realidad (*¡¡Desconocida!!*) H_0 Cierta	H_0 Falsa
H_0 Rechazada H_1 Aceptada	Error tipo I (α) Falsos (+)	Decisión correcta
H_0 NO Rechazada	Decisión correcta	Error tipo II (β) Falsos (-)

Error alfa (α) = Probabilidad de equivocarnos si rechazamos la hipótesis nula (Ho) cuando ésta es cierta.

Error beta (β)= Probabilidad de equivocarnos si NO rechazamos la hipótesis nula, a pesar de que sea falsa (H_1 cierta).

Potencia del test (1-β)= probabilidad de rechazar la hipótesis nula cuando es falsa (encontrar diferencias cuando éstas existen)

Tabla II.- Esquema simplificado de elección del test de contraste de hipótesis

		Variable Dependiente	
Variable Independiente	**Nominal**	**Ordinal (continuas no normales)**	**Continua (razón o intervalos)**
Nominal Dicotómica (2 muestras)	***Muestras Independientes:*** - Test Z comparación de proporciones - Test Ji cuadrado - Test Exacto Fisher	Test U Mann Whitney (Wilcoxon suma rangos)	Test t Student muestras independientes
	Muestras Relacionadas: Test McNemar Test Z y Método Binomial	T.Wilcoxon rangos con signo	Test t Student muestras apareadas
Nominal Politómica (> 2 muestras)	Test Ji cuadrado Método binomial	Test de Kruskal - Wallis * Muestras apareadas: Prueba de Friedman	ANOVA
Continua	Test t Student	Coeficiente Correlación de Spearman	Coeficiente Correlación de Pearson

Tabla III.- Análisis de resultados seleccionados del estudio sobre abuso del uso de móvil, obesidad y problemas de sueño.

	Abuso de móvil en semana					
	No n=101		Si n=113		p	Prueba
Sexo masculino n (%)	52	(51,5%)	46	(40,7%)	0,114	Ji cuadrado
Obesidad n (%)	18	(18,0%)	34	(30,4%)	**0,037**	Ji cuadrado
Estudios Madre n (%)					**<0,001**	Ji cuadrado
Primaria	37	(36,6%)	61	(54,5%)		
Secundaria	21	(20,8%)	31	(27,7%)		
Bachiller/FP/Superior	43	(42,6%)	20	(17,9%)		
Talla (Zscore) X (DE)	0,28	(1,11)	0,20	(0,90)	0,772	t Student
Demora acostarse (minutos) M (RIC)	-30	(-60; 0)	0	(-30; 60)	**<0,001**	Mann-Whitney

	IMC (Zscore)			
	X	DE	**p**	Prueba
Sexo			0,236	t Student
Masculino n=98	0,79	1,70		
Femenino n=116	1,10	2,03		
Estudios Madre			**0,036**	ANOVA
Primaria	1,21	2,06		
Secundaria	1,13	1,84		
Bachiller/FP/Superior	0,45	1,55		

DE: desviación estándar; Demora acostarse: demora respecto las 23:00 horas; FP: formación profesional; M: mediana; n: recuento; RIC: rango intercuartílico (percentil 25; 75); X media; Zscore: medida estandarizada por edad y sexo.

Figura 1.- Histograma de la variable talla estandarizada (Distribución normal)

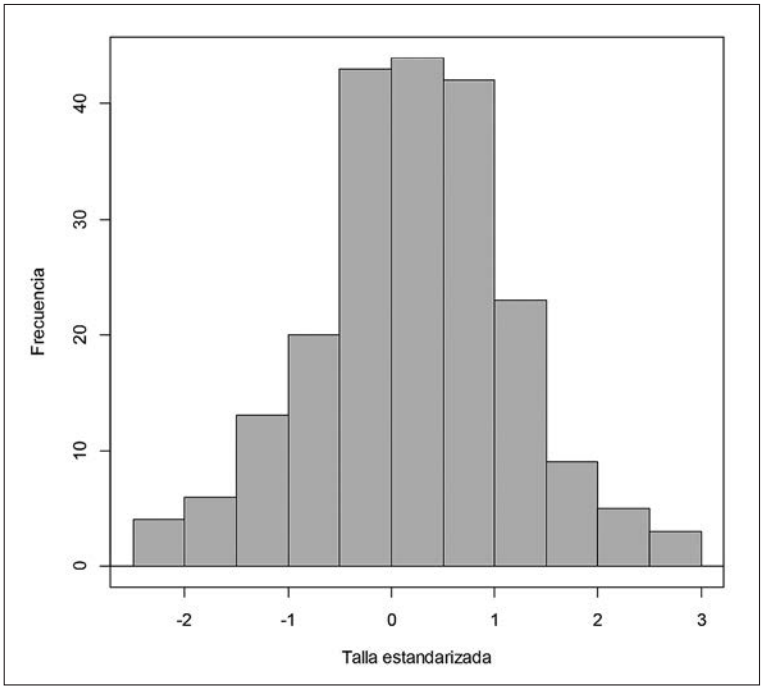


Figura 2.-Distribución de probabilidad normal.

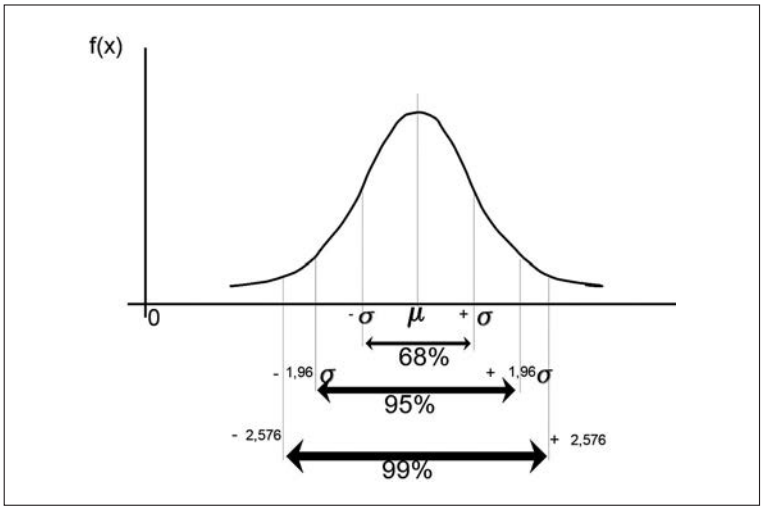
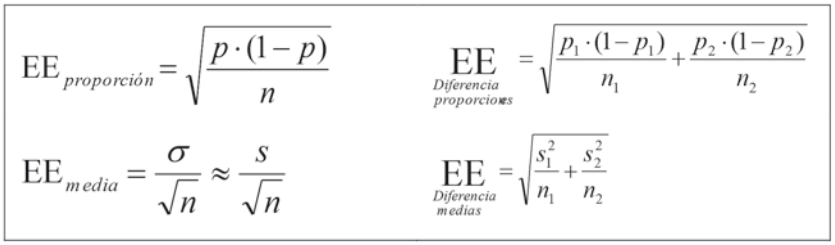


Figura 3.- Fórmulas de estimación de errores estándar de algunos parámetros poblacionales:



EE error estándar; **n** tamaño muestral; **p** proporción; **s** desviación típica muestral; σ desviación típica poblacional