

## Errores estadísticos

Florencia Reckziegel<sup>1</sup>

<sup>1</sup> Instituto de Bio y Geociencias del NOA (IBIGEO, UNSa – CONICET). e-mail: florenciareckziegel@gmail.com

La estadística es una rama de la matemática de extendido uso. La utilizan, por ejemplo, quienes nos dan cifras de predicciones para próximas elecciones, o de cuál será el valor del dólar la próxima semana, o para obtener índices de pobreza, de crecimiento de poblaciones, etc. Como no podemos tener acceso a la información de toda una población, la estadística se encarga de trabajar sobre muestras y establecer con ellas, conclusiones sobre la población. Con valores de errores debidos a la naturaleza del azar y de desconocer la información total. Si la muestra realmente representa a la población, pueden existir diferentes factores, como la aleatoriedad que conduzca a unas medidas finales de la muestra que se separen un poco de las medidas poblacionales. Con más razón nos equivocaremos en las conclusiones cuando se toma una muestra que no representa realmente a la población. Si queremos inferir, por ejemplo, sobre la religión que más adeptos tiene en un país, estaría muy mal que tomemos una muestra conformada por personas de una única religión. O si queremos saber la opinión sobre la política del país, estaría mal preguntar sólo a nuestros amigos. En esos casos, no es la estadística la que se equivoca, sino la persona que está dirigiendo un estudio sesgado y la usa de manera incorrecta.

Sobre esto les vengo a hablar, porque en estadística existen los errores inherentes a los métodos utilizados. Las conclusiones en estadística se refieren (o se deberían referir) con un intervalo de confianza y un margen de error que se acepta, se conoce y se debe informar. Pero también existen los errores humanos, evitables, involuntarios (o quizás no siempre) que conducen a conclusiones o interpretaciones erróneas. Conocer cómo funciona y como evitar estos últimos errores, nos permitirá notarlos para no ser engañados y para no repetir errores cuando usemos esta herramienta que sí es muy útil cuando se la usa bien.

La matemática (y la estadística) nos proporciona herramientas y métodos para abordar problemas en otras áreas científicas:

1. Son esenciales para comprender los procesos naturales y modelarlos. Por ejemplo, en la geología, las ecuaciones matemáticas y los modelos estadísticos se utilizan para predecir la distribución de recursos naturales como el petróleo y el gas, la ubicación de depósitos minerales y la predicción de tsunamis o dispersión de ceniza volcánica, dada una erupción. En biología, se utilizan para modelar la evolución de las especies, la dinámica de las poblaciones, la respuesta de los organismos a los cambios ambientales, los efectos de los cambios climáticos, entre muchos otros.

2. Se utilizan para analizar grandes conjuntos de datos, lo que permite a los científicos extraer patrones y relaciones que de otra manera serían difíciles o imposibles de detectar. Los análisis estadísticos también permiten a los científicos evaluar la significancia de los resultados y determinar si los hallazgos son el resultado del azar o si realmente hay una correlación o causa y efecto o una tendencia.

3. Son esenciales para la toma de decisiones. Por ejemplo, en la geología, los modelos matemáticos se utilizan para determinar la ubicación óptima de pozos de petróleo y gas, mientras que, en la biología, pueden ayudar a predecir el impacto de la introducción de una nueva especie en un ecosistema.

Sin embargo, la mala aplicación de estas herramientas puede comprometer la validez de los resultados obtenidos y, por lo tanto, la confiabilidad de la investigación. En este artículo, se expondrán algunos de los errores estadísticos más comunes que se observan en la investigación científica. Se discutirá la importancia de utilizar la matemática de manera adecuada. Se explicará por qué ocurren estos errores y se ofrecerán sugerencias prácticas para evitarlos.

## Casos particulares

### *-Correlaciones no siempre relacionadas*

Decir que existe una correlación entre dos variables es afirmar que los cambios de una variable están relacionados con los cambios en la otra variable.

Sin embargo, se suele caer en un error cuando se asume que una correlación entre dos variables implica necesariamente una relación causal entre ellas, cuando en realidad puede ser una coincidencia o estar, estas dos variables, influenciadas por otra/s variables no consideradas.

Un ejemplo de esto, es la correlación observada entre la cantidad de terremotos y la actividad humana en una región determinada, si se evalúan las dos variables y ambas están tomando valores que

van en aumento a lo largo del tiempo, algunas personas podrían asumir que la actividad humana está causando los terremotos, cuando en realidad, el aumento de actividad sísmica es inherente a factores geológicos naturales y puede coincidir con regiones donde se ve un aumento de la población y por ende, sus actividades normales nada tienen que ver con los sismos. Con lo cual, la correlación de los dos aumentos de actividad es en realidad, una mera coincidencia.

Otro ejemplo de que correlación no implica causa y efecto es la relación entre la masa corporal de mamíferos y la latitud. En este caso, se observa una correlación positiva en la que, a menores latitudes, la masa corporal tiende a ser menor, mientras que, a mayores latitudes la masa corporal tiende a ser mayor. Sin embargo, es importante destacar que la latitud en sí misma no causa directamente el aumento de la masa corporal de los mamíferos. En realidad, la temperatura es la variable determinante en esta relación, ya que a medida que nos desplazamos hacia latitudes más altas, la temperatura disminuye, lo que influye en la masa corporal de los mamíferos. Por lo tanto, la temperatura es el factor subyacente que explica esta correlación, en lugar de la latitud en sí misma.

Es importante tener en cuenta que las correlaciones pueden ser útiles para identificar patrones y relaciones potenciales entre variables, pero no siempre indican una relación causal. Es necesario realizar estudios más detallados y cuidadosos para determinar la verdadera naturaleza de la relación entre las variables en cuestión.

*- Porcentajes, test y enfoques, no nos dejemos engañar*

Cox Paul (1998) redactó un “glosario de errores matemáticos”, con ejemplos muy prácticos de algunos errores típicos que se cometen, entre los cuales mencionan la conjetura del falso positivo. La cual te invito a analizar: se trata de un test con una precisión muy alta, pero no perfecta que puede producir más falsos positivos que verdaderos positivos.

Y acá hago una pausa para explicarte qué es la precisión de un test y de qué se habla cuando se dice falsos y verdaderos positivos (también existen los falsos y verdaderos negativos).

La precisión de un test se puede expresar en términos de porcentaje, y se refiere a la proporción de resultados correctos que un test da en relación al total de resultados obtenidos. Por ejemplo, una precisión del 90% significa que, de cada 100 resultados obtenidos por el test, 90 son correctos y 10 son incorrectos. Es decir, el test tiene una alta probabilidad de dar resultados exactos y confiables en un 90% de los casos.

Entonces, volviendo a la conjetura, decíamos que un test puede conducir a más falsos positivos que verdaderos positivos si la precisión es muy alta (pero no perfecta).

Por ejemplo: Supongamos que un 3% de una población hace uso de una droga ilegal, X. Supongamos también que se desarrolló un test, con un 95% de precisión, que permite determinar si un individuo ha estado usando esa droga X. Ahora, supongamos que se aplica este test a 1000 personas (Figura 1), de las cuales 30 son usuarios de X (el 3%). Como el test tiene una precisión del 95%, 29 de los usuarios serán descubiertos (el otro 5% obtiene un falso negativo: es decir el test le da negativo cuando en realidad es un usuario). Al mismo tiempo, de las 970 personas restantes, 48 de ellas (el 5%) también dan positivo, aunque en realidad no sean usuarios (estos son los falsos positivos). En otras palabras, 78 personas dieron positivo para el uso de la droga X, pero solo 29 fueron verdaderos positivos. Con lo cual se obtuvieron más falsos positivos que verdaderos positivos.



Figura 1: Representación gráfica de la conjetura del falso positivo.

Podes intentar sacar tus propias cuentas para el test del HIV que tiene un 99,7% de precisión, o para el polígrafo (detector de mentiras) que tiene una precisión de sólo un 80%.

Un aporte final: Es importante tener en cuenta que la precisión no lo es todo. También es necesario evaluar otros factores como la sensibilidad y especificidad del test, que se refieren a la capacidad del test de detectar verdaderos positivos y negativos, respectivamente. Además, la precisión de un test puede variar según la población evaluada y otros factores como la calidad de las muestras o la experiencia del personal que realiza la prueba.

Otro ejemplo interesante que menciona Paul Cox es la paradoja de Simpson: Esta paradoja trata sobre una condición existente en varios grupos de datos, que desaparece cuando se agrupan todos esos datos, más aún: la tendencia con los datos agrupados es contraria a la de los grupos individuales. Esto se produce cuando se establecen conclusiones (en los grupos) sin considerar otras variables presentes.

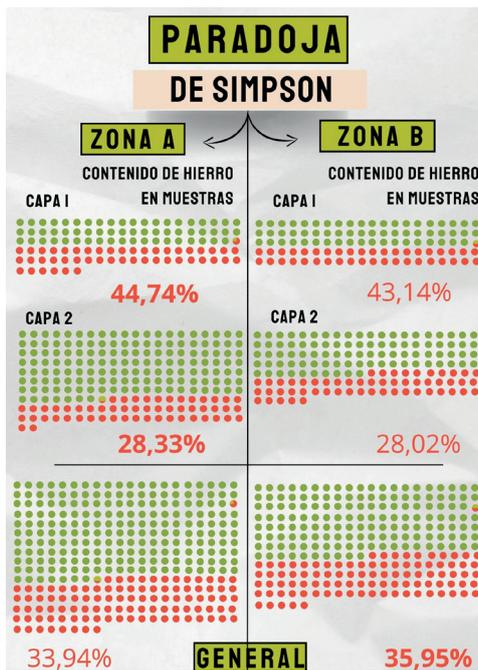


Figura 2: Representación gráfica de la paradoja de Simpson.

Un ejemplo para ilustrar esta paradoja (obtenido en Trujillo et al., 2012 y representado en la figura 2) trata sobre el contenido de hierro en rocas. Se analizan dos zonas y, en cada zona, dos niveles: capa 1 y capa 2. Cuando se analizan los niveles por separado, se observa que la zona A tiene un contenido de hierro superior que la zona B. Sin embargo, al analizar los dos niveles en conjunto resulta que la zona B tiene una mayor proporción de hierro.

Dependiendo de cómo se agrupen los datos, el porcentaje de hierro parece más favorable en la zona A (donde cada capa es más rica en proporción, que su contraparte en la zona B), o si se considera al general (ambas capas), la zona B es más rica en proporción que la zona A. Si usted necesitara vender una zona de éstas y

el contenido de hierro fuera un factor importante ¿qué agrupamiento mostraría?

Mostrar uno u otro según un interés personal es un ejemplo de la manipulación de la estadística con el único objetivo de engañar a la otra parte. Por ética, no es correcto manipular hechos o situaciones para obtener números que nos permitan vender un producto o convencer a alguien sobre una idea. A nuestro alcance tenemos el poder de aprender para no dejarnos engañar.

La lista de Paul Cox suma muchos casos y ejemplos más. Muy interesantes, típicos y muy reales.

### Errores matemáticos en la investigación científica

En muchos trabajos científicos, recurrir a pruebas estadísticas es clave, aunque a veces su aplicación carece de la rigurosidad necesaria, dando lugar a interpretaciones erróneas o resultados inexactos. Supongamos que pretendemos conocer el tamaño medio de las partículas de una muestra de arena de una zona específica. Nunca vamos a poder medir todas las partículas, se miden algunas y a partir de ellas se calcula su tamaño medio. Probablemente no sea igual al verdadero tamaño medio de la población de partículas de arena, pero seguramente se aproximará. Las distintas estimaciones podrán ser diferentes, pero se distribuirán alrededor del verdadero valor medio. Para cada estimación podemos establecer un intervalo, llamado intervalo de confianza, dentro del cual se confía que esté el verdadero valor medio. En la figura 3, se observa un intervalo de confianza del 95%, esto significa que, si tomamos muchas muestras diferentes, medimos los tamaños de arena y

generamos esos intervalos de confianza, en el 95% de ellos se incluirá en la zona verde al valor medio de la población.

Supongamos que tenemos una estimación hecha en una muestra que no estamos seguros si el tamaño se corresponde al de la población en cuestión. Teniendo el valor estimado del tamaño medio de esta muestra, podemos compararlo con el de la población haciendo una prueba estadística. Nos preguntamos si esta muestra es parte de la población. Si el valor medio medido en la muestra es de 0,25 mm podría serlo (ver Figura 3); si fuera de 0,01 o de 0,8 mm estaremos bastante seguros de que no pertenece a la población. Un valor medio de 0,3 mm, aun siendo algo distinto (de 0,25 mm) podría ser parte de la misma población, y la diferencia de los valores medios ser debida al azar y/o a los errores del muestreo.

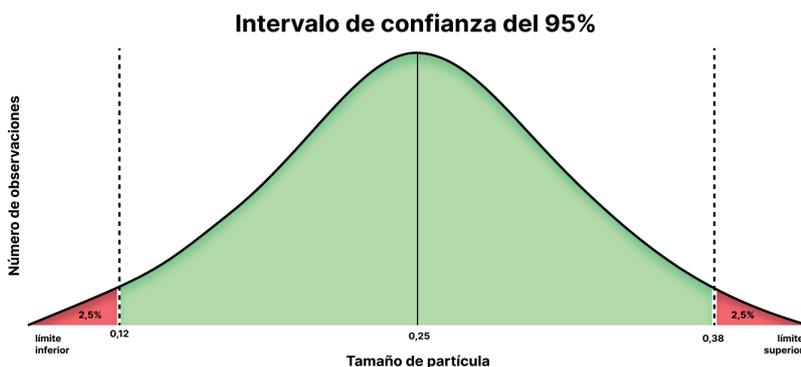


Figura 3: Intervalo de confianza para la media. Tenemos una confianza del 95% de que el intervalo marcado en verde, contiene al valor de la media poblacional. La línea recta vertical indica el valor medio. Las líneas punteadas dividen la zona verde: en la que se espera que esté la media de la población a la que pertenece la muestra; de la zona roja: en la que se rechaza la hipótesis nula que suponía la igualdad de las medias de la muestra y de la población.

La prueba estadística requiere definir una hipótesis nula ( $H_0$ ) y un estadístico de prueba. La hipótesis nula en este caso sería que el valor medio medido en la muestra coincide con el valor medio de la población. No tiene por qué coincidir exactamente, pero existe la posibilidad de que la diferencia esté dentro del margen permitido por el azar. Crearemos un estadístico de prueba basado en lo que sabemos de la población, que permitirá calcular cuáles son los límites de tamaño, por encima y/o por debajo de los cuales sería casi imposible que la diferencia se debiera al azar. Y calcularemos un valor de  $p$  para nuestro dato: la probabilidad de que la diferencia entre el valor obtenido y el de la población sea debido al azar. Consideramos falsa a la hipótesis nula de que los dos valores son iguales o solo difieren por azar, si la probabilidad  $p$  es menor que un nivel de significación establecido (por lo general menor que 0,05).

Teniendo en cuenta estos conceptos, veamos una lista de errores matemáticos en los que se suele incurrir en la investigación:

- *Mala presentación de los números:* Dar cifras decimales sin sentido que confunden sin agregar información. Por ejemplo, reportar el peso de un paciente como 60,18 kg, en lugar de 60 kg (dependiendo del estudio que se realice). Lo mismo puede suceder con un valor de probabilidad y otras cifras decimales. Asimismo, dar valores en diferentes unidades de medida genera confusión, es importante unificar estos puntos a lo largo de una publicación.

- *Dividir datos continuos en categorías sin explicar por qué o cómo:* Al hacerlo se puede reducir la precisión de las mediciones y la variabilidad de los datos conduciendo a errores por pérdida de información. A menos que esto esté bien justificado y se incluya la justificación.

- *Una presentación inapropiada de los datos:* tal como usar estadística descriptiva de manera incorrecta. Es habitual utilizar medidas propias de una distribución normal, como la media y la desviación estándar, para describir datos que no tienen esa distribución. Cuando los datos siguen una forma normal significa que se distribuyen de manera simétrica alrededor del promedio dispersándose más o menos de acuerdo a una medida de dispersión de datos alrededor de la media que es el desvío estándar (Figura 4A). Luego, si los datos no se distribuyen de manera normal (Figura 4B), no tiene el mismo sentido reportar la media y el desvío. En esos casos se recomiendan medidas como la mediana y el rango.

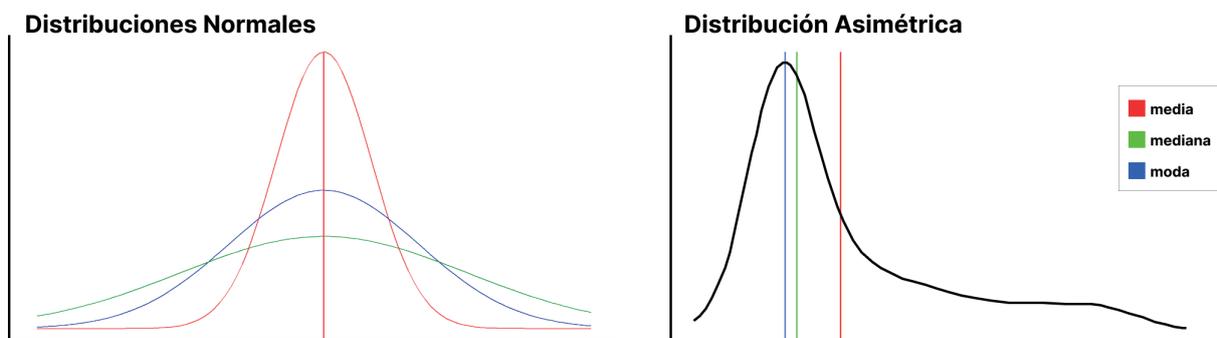


Figura 4: Izquierda: Cada una de las tres curvas corresponde a distribuciones normales que tienen el mismo valor medio (línea vertical roja) pero distintas dispersiones. La dispersión (el "ancho" de la curva) se cuantifica con el valor de desvío estándar. Las tres son distribuciones normales porque se distribuyen de igual manera hacia los dos lados de la media. Derecha: Distribución no normal. Para las mediciones que tienen esta forma ocurre que la media (línea vertical roja) no indica que los datos se dispersan de igual manera hacia ambos lados. Por lo tanto, referirse a la media para caracterizar esta distribución no es suficiente.

<sup>1</sup>Los datos continuos son aquellos que se pueden medir y representar en una escala numérica, tal que pueden tomar cualquier valor dentro de un rango, sin saltos e interrupciones, incluyendo valores infinitamente precisos en forma de números decimales.

- *Mala interpretación del valor  $p$  obtenido en pruebas estadísticas:* Existen muchos enunciados errados acerca del valor de probabilidad  $p$ , el que permite rechazar o no una hipótesis nula cuando se realiza un test estadístico. Es incorrecto, por ejemplo, decir que “el valor  $p$  es la probabilidad de que la hipótesis nula ( $H_0$ ) sea verdadera”. Lo cierto es que  $p$  es la probabilidad de obtener por azar un valor tan alejado del verdadero como el obtenido.

- *Realizar pruebas sólo para tener un valor  $p$ .* La significancia de un test (cuando  $p$  es menor que el valor determinado, por lo general 0,05) sirve para tomar decisiones. Pero no siempre hay que tomar decisiones, lo cual significa que no siempre hay que aplicar una prueba de hipótesis (como cuando queremos explorar o describir un conjunto de datos).

- *Significancia de un test (un valor  $p$  suficientemente pequeño para rechazar la hipótesis nula) no es igual al tamaño del efecto del test:* Si el resultado de un test es significativo y logramos rechazar la  $H_0$ , esto no supone que el tamaño del efecto sea grande. La conclusión de la significancia de un test debe aplicarse sobre la fortaleza de la evidencia y no sobre el tamaño del efecto. Así como también es importante saber que no encontrar una evidencia significativa no significa que no haya o que no exista. Cuando el tamaño de la muestra es grande, diferencias pequeñas pueden ser estadísticamente significativas, pero eso no significa que sean científica, biológica, geológica o clínicamente significativas.

- *Diferencia no significativa no es lo mismo que un efecto ausente:* Mientras que algunos cometen el error de suponer que, si un  $p$  es suficientemente grande como para no poder decir que la prueba es estadísticamente significativa, implica que la  $H_0$  es cierta. Cuando en realidad, sólo implica que no tenemos suficiente evidencia para rechazar la  $H_0$ . Por eso se insiste en que, si  $p$  no toma valores inferiores al nivel de significación deseado, no es que se acepte la  $H_0$ , si no que no se la rechaza. A veces se tiende a interpretar estudios con resultados no significativos y baja potencia estadística como “negativos”, cuando en realidad son no concluyentes.

- *No confirmar que los datos cumplan con los supuestos de las pruebas estadísticas utilizadas para analizarlos:* lo que puede dar lugar a resultados inexactos. Se debe incluir tanto el nombre de la prueba como una declaración de que sus supuestos se cumplieron en cada análisis estadístico informado. Usar un método estadístico sin saber si los datos cumplen los supuestos del método o en algunos casos incluso sabiendo que no cumplen los supuestos conduce a conclusiones indemostrables y sin sustento. Un método estadístico está probado que funciona y que arribe a conclusiones determinadas bajo condiciones determinadas. Si las condiciones no se verifican, nada de todo lo que sigue (ni

procedimientos, ni resultados), tendrá sentido ni puede tomarse como cierto.

- *Tamaño de muestra inapropiado*: Cuando se realiza una prueba de hipótesis sobre muestras muy pequeñas o sesgadas, se pueden cometer errores (inherentes al método) al sacar conclusiones:

1. Error de falso positivo: rechazar la hipótesis nula cuando ésta es verdadera. Esto sucede cuando las medidas que obtenemos de la muestra nos dan un valor casualmente (debido al azar) lejano del valor poblacional. Existe una probabilidad muy chica (por lo general se escoge menor a 0,05) de que esto suceda.

2. Cuando el tamaño de la muestra es muy pequeño, hay mayor probabilidad de cometer un error por falso negativo, lo cual significa que la prueba tiene menor poder estadístico. Esto se da cuando no rechazamos la hipótesis nula y esta sí es falsa. Cuando la variabilidad de la muestra obtenida no es suficiente y obtenemos medidas que no nos permiten detectar la diferencia con respecto a la medida poblacional, aun cuando esa diferencia sí existe. Revisar el ejemplo mencionado antes sobre la droga X (que puede extenderse a test de covid u otros).

- *Eliminación o falta de datos*: La omisión o manipulación de datos representa una problemática recurrente, donde en ocasiones se excluyen datos atípicos sin justificación, pudiendo contener información valiosa. Además, la eliminación de datos con valores faltantes, un muestreo inadecuado, o la focalización en sub-rangos en lugar del rango completo en estudios que abordan características de amplios grupos demográficos, pueden comprometer la integridad de los resultados.

- *Tablas y figuras que no comunican la realidad*. En presentaciones científicas, las tablas y figuras deben usarse para comunicar información y no simplemente para almacenar datos y es importante seleccionar la forma adecuada para presentar los datos para alentar a los lectores a hacer una comparación específica. Es un acto de mala conducta y una falta de ética usar un gráfico o tabla en el que el mensaje visual no respalda el mensaje de los datos en los que se basa, lo que puede ser engañoso y distorsionar la interpretación de los datos presentados.

### ¿Por qué ocurren?

Los errores en la investigación científica referente a estadística pueden deberse (no excluyentemente) a:

- *Ignorancia*: Se deben a la falta de conocimiento o comprensión adecuados del problema de investigación o del método utilizado.

- *Planificación de estudio deficiente*: Una mala planificación del estudio, como la selección inadecuada de los participantes (de la muestra), la falta de control de variables relevantes o una muestra demasiado pequeña, puede llevar a errores en la investigación científica.

- *Expectativas de publicación*: La presión para publicar en revistas científicas de alto impacto puede llevar a la manipulación de los resultados o a exagerar las conclusiones del estudio, lo que puede dar lugar a errores en la investigación científica.

- *Emoción*: Los investigadores también pueden verse afectados por sus propias emociones, como la frustración, la ansiedad o el entusiasmo excesivo, lo que puede llevar a errores en la interpretación de los datos o la selección de los métodos de análisis.

- *Recursos*: Para llevar a cabo una investigación científica rigurosa, se necesitan recursos adecuados como tiempo, personal, educación o dinero, y la falta de ellos puede conducir a errores en el estudio.

- *Prioridades en conflicto*: Los investigadores pueden tener prioridades en conflicto que compiten por recursos, atención o fuerza de voluntad, lo que puede resultar en ciencia descuidada, comportamiento negligente o distorsión de las observaciones. Esto puede deberse a una variedad de factores, desde creencias personales hasta prejuicios de la industria y tendencias culturales.

### Cómo prevenirlos

Existen varias medidas que se pueden tomar para evitar errores matemáticos en las ciencias. Te presento algunas recomendaciones:

- *Mejorar las prácticas del diseño de estudio*: Es importante planificar los análisis estadísticos de antemano para que todo comience con una correcta toma de muestras. Considerar buenas prácticas de calidad para la recopilación y gestión de datos.

- *Utilizar estadísticas y modelos adecuados*: Es importante elegir el modelo estadístico adecuado para los datos y la pregunta de investigación en cuestión (qué tengo, qué quiero), y asegurarse de que los supuestos del modelo se cumplan (con este modelo o método estadístico ¿puedo responder la pregunta? ¿Se cumplen todos los supuestos y puedo realmente usarlo?). También es importante

verificar la calidad de los datos y, si es necesario, eliminar los valores atípicos o los datos que son claramente erróneos (justificando, siempre actuar en función de lo correcto).

- *Validar los modelos y las hipótesis:* Los modelos y las hipótesis deben ser validados mediante pruebas rigurosas y, si es posible, mediante experimentos independientes y repeticiones. También es importante realizar análisis de sensibilidad para determinar la robustez de los resultados a los cambios en los parámetros y las condiciones del modelo. Utilizar validaciones cruzadas y muestras externas siempre que corresponda para confirmar los resultados.

- *Considerar múltiples factores:* Es importante considerar múltiples factores que puedan contribuir a los resultados observados, y controlar o eliminar los efectos de los factores confusos. Por ejemplo, en un estudio sobre los efectos de la contaminación del aire en la salud humana, es importante controlar los efectos de otros factores como el tabaquismo o la dieta.

- *Divulgación y revisión:* Los resultados deben ser divulgados de manera clara y transparente, y los métodos y los datos deben estar disponibles para la revisión y la verificación por parte de otros investigadores (esto lo hace fiable y reproducible). También es importante que los resultados sean revisados por otros expertos en el campo antes de ser publicados.

- *Formación y colaboración:* Es importante que los investigadores y revisores reciban una formación adecuada en estadística y métodos matemáticos, y que trabajen en colaboración con expertos en estas áreas. Además, es importante mantenerse actualizado sobre los avances en estas áreas y aplicar nuevas técnicas y herramientas cuando sea necesario.

*“La verdadera solución para una mala presentación estadística se dará cuando los autores aprendan más sobre el diseño de investigación y las estadísticas; los estadísticos mejoren su capacidad para comunicar estadísticas a los autores, editores y lectores; los investigadores involucren a los estadísticos al comienzo de la investigación; los editores de manuscritos comprendan y apliquen las directrices de presentación estadística, cuando más revistas puedan examinar con más cuidado los artículos que contengan análisis estadísticos; y los lectores aprendan más sobre cómo interpretar las estadísticas y comiencen a esperar, si no exigir, una presentación estadística adecuada (Tom Lang, 2004).”*

---

#### REFERENCIAS Y LITERATURA RECOMENDADA

---

BROWN AW, KA KAISER, DB ALLISON. 2018. Issues with data and analyses: Errors, underlying themes, and potential solutions. *Proceedings of the National Academy of Sciences*, 115 (11): 2563-2570.

COX P. 1998. Glossary of mathematical mistakes. <http://www.mathmistakes.com>

LANG T. 2004. Twenty statistical errors even you can find in biomedical research articles. *Croatian Medical Journal*, 45(4): 361-370.

MATAMOROS PINEL RA, A CEBALLOS MÁRQUEZ. 2017. Errores conceptuales de estadística más comunes en publicaciones científicas. *CES Medicina Veterinaria y Zootecnia*, 12 (3): 211-229.

NUIJTEN MB. 2016. Preventing statistical errors in scientific journals. *European Science Editing*, 42 (1): 8-10.

NUZZO R. 2014. Statistical errors. *Nature*, 506 (7487): 150.

TRUJILLO CODORNÍU RA, S BERNAL-HERNÁNDEZ, M RASÚA LÓPEZ. 2012. La paradoja de Simpson en la exploración de yacimientos lateríticos cubanos. *Minería y Geología*, 28 (2): 1-12.

WEINBERG CR, D ZAYKIN. 2015. Is bad luck the main cause of cancer? *Journal of the National Cancer Institute*, 107 (7): 1-4.