



## ARTÍCULO ESPECIAL

# Estadística en el siglo XXI

*Statistics in the 21st Century*

Héctor HEVIA <sup>a</sup>✉

<sup>a</sup> Universidad Adolfo Ibáñez, Chile  
 Universidad Alberto Hurtado, Chile

### INFORMACIÓN DEL ARTÍCULO

#### Historia del Artículo:

Recibido: 01 01 2025  
 Aceptado: 01 04 2025

#### Palabras claves:

Estadística. Pensamiento Estadístico. Matrices de Koestler. Bisociaciones. Interpretación. Significado.

#### Key words:

Statistics. Statistical Thinking. Koestler Matrices. Bisociations. Interpretation. Meaning.

### RESUMEN

En toda situación de interés que se estudie, donde sea posible extraer datos, se confía que habrá procedimientos estadísticos que evidenciarán parte del conocimiento de valor que encierran tales datos. La universalidad de la Estadística es innegable; como es innegable su valor y su belleza. Durante años, la Estadística se consideró como parte de la Matemática; pero, un análisis serio permite rechazar esta idea (ver Moore, 1988). Sorprendentemente, su íntima relación con el razonamiento inductivo tampoco fue suficientemente enfatizada; Moore señaló insistentemente esta falencia, y un somero análisis resulta evidente. Se sostiene que el aprendizaje de la Estadística debe ligarse íntimamente a la programación; olvidando que el aprendizaje en el ser humano, no sufrió mayor cambio, pese al notable avance de las TIC; la tecnología debe estar al servicio del aprendizaje del ser humano (ver Bruner, 1964). La Estadística requiere una revisión en su pensamiento, pero, ¿cómo es posible hacer una revisión de un pensamiento aún no depurado y sometido a más de un paradigma como está señalado? La dificultad que se presenta es que la indagación en un cierto pensamiento debe realizarse en el mismo acto en que este pensamiento se realiza, como una experiencia de la conciencia. Frente a esto, una alternativa que surge, es que esta indagación se realice guiada por el conocimiento y métodos que la Fenomenología de Husserl provee. En este artículo busca establecer coherentemente y desde una perspectiva fenomenológica, un escenario de certezas que permitan el desarrollo natural del pensamiento estadístico, para cualquier individuo.

### ABSTRACT

In any situation of interest that is studied, where it is possible to extract data, it is trusted that there will be statistical procedures that will demonstrate part of the valuable knowledge contained in such data. The universality of Statistics is undeniable; as its value and beauty are undeniable. For years, Statistics was considered part of Mathematics; but, a serious analysis allows us to reject this idea (see Moore, 1988). Surprisingly, its intimate relationship with inductive reasoning was also not sufficiently emphasized; Moore insistently pointed out this shortcoming, and a cursory analysis is evident. It is argued that the learning of Statistics should be closely linked to programming; forgetting that human learning did not undergo major changes, despite the notable advance of ICT; Technology must be at the service of human learning (see Bruner, 1964). Statistics requires a review of its thinking, but how is it possible to review a thought that has not yet been refined and is subject to more than one paradigm as indicated? The difficulty that arises is that the investigation into a certain thought must be carried out in the same act in which this thought is realized, as an experience of consciousness. Faced with this, an alternative that arises is that this investigation is carried out guided by the knowledge and methods that Husserl's Phenomenology provides. This article seeks to establish coherently and from a phenomenological perspective, a scenario of certainties that allow the natural development of statistical thinking, for any individual.

#### ✉ Autor para correspondencia

Correo electrónico: [hctr.hevia@gmail.com](mailto:hctr.hevia@gmail.com)

<https://doi.org/10.63706/jsibemir.v1i1.6>

e-ISSN: 3087-2367/© 2025 JS

Este es un artículo Open Access bajo licencia BY-NC-ND  
 (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## INTRODUCCIÓN

Se reconoce el origen de la Estadística, y, por tanto, del pensamiento estadístico, en el trabajo con datos según D. S. Moore: “El tema de la estadística son los datos y la inferencia a partir de datos”, D. S. Moore; quien más adelante señala:

*Los orígenes de la estadística yacen en la recopilación de datos oficiales y privados, en los censos y padrones y tablas de mortalidad. La fundación tanto de la Sociedad de Estadística de Londres (1834) y de la Asociación Estadounidense de Estadística (1839) se debió a la preocupación por la adecuada recopilación y tabulación de datos. Los métodos estadísticos modernos para la inferencia de datos surgen de problemas prácticos –primero en astronomía y agrimensura, y luego en las ciencias sociales y de la vida. El mayor genio de la estadística del siglo veinte, R. A. Fisher, se preocupó por diseñar y analizar experimentos en agricultura y fue nombrado caballero por su trabajo en genética. La valiosa claridad y precisión aportadas por los métodos matemáticos, especialmente durante la generación de la postguerra, no han cambiado el hecho de que **la estadística deriva su importancia y su misma razón de ser, de sus raíces en los datos y en la inferencia científica.** [2]*

No se trata solo de datos separados de su contexto, sino, por el contrario, de **datos en contexto**. Y el paradigma fundamental en esta concepción es que **el contexto no puede ser sustituido, tampoco simulado**; al menos, no en forma absoluta. Ello explica la permanente necesidad de la obtención de datos, ya sea bajo condiciones naturales propias del contexto, o experimentales bajo el control de ciertas variables, con el objetivo que, a través de la interpretación de estos datos ya procesados, se agregue nuevo conocimiento para una toma de decisiones basada en información relevante. Por ende, resulta obvio que esta interpretación deba darse en el contexto desde donde estos datos fueron originados. Observemos, en este punto, que el contexto de los datos no es la población, por muy amplio que sea el significado que se le otorgue al concepto “población”. **El contexto es la realidad que subyace**, como señala Moore: “La estadística se preocupa de la recopilación, organización y análisis de los datos, y de las inferencias desde los datos a la realidad subyacente” [2]. Esta realidad subyacente es un universo de significados; son múltiples universos de pensamiento para quienes participan de esta realidad, la que juega un rol crucial en la interpretación. A continuación, desarrollamos estas ideas siguiendo a Arthur Koestler [3]. Como introducción a las nociones fundamentales de la fenomenología de Husserl, ver Capítulo 2 de D. Langdrige, 2007 [4].

## INTERPRETACIÓN EN ESTADÍSTICA Y MATRICES DE KOESTLER

Al parecer, habría sido Joseph Tal en su libro *Reading between the numbers, Statistical thinking in everyday life*, quien señaló, por primera vez, la relevancia de las denominadas “matrices de pensamiento de Koestler” como marco conceptual para entender cómo toma lugar la interpretación de los resultados de un cierto procedimiento estadístico, que ha sido aplicado a los datos de interés [5].

*La explicación estadística ocurre cuando, para usar las palabras de Koestler, ‘dos matrices diferentes’ de fenómenos se conectan. Es salir de un fenómeno particular para comprenderlo. (...) Esta relación de diferentes contextos entre sí es la esencia de la estadística inferencial.* [5]

Revisando, en primer lugar, estas “matrices” a las que alude J. Tal, recordando que para Koestler el sentido de “matriz” es “molde”, “entidad principal generadora de otras”; una matriz de Koestler es un universo de discurso, un contexto que se extiende en un horizonte donde prima la coherencia que entrega el sentido común [3][5]. Otros términos utilizados por Koestler para denomi-

nar a estos universos de discurso son: marcos de referencia, contextos asociativos, tipos de lógica, códigos de comportamiento; más técnicamente, matrices de pensamiento (o matrices de comportamiento) [3]. Al respecto, Hevia señala:

*Arthur Koestler (1964) estudia tres dominios de la creatividad: humor, ciencias y arte, buscando aspectos comunes a los actos de creación que se desarrollan en cada uno de estos tres ámbitos. Consideraciones sobre el primer dominio, permiten al autor establecer que en una situación de humor se hace presente el choque de dos contextos cuyas lógicas o códigos de comportamiento son incompatibles. En estos casos, se percibe una situación o idea L en dos marcos de referencia, M1 y M2; cada uno de estos marcos consistente en sí mismo, pero, incompatibles entre sí, los cuales son actualizados en la conciencia en forma simultánea ‘el evento L vibra en dos diferentes longitudes de onda, como tal’ (Koestler, p 35). La situación L, no está únicamente vinculada a cada contexto por separado, sino asociada a ellos de una manera singular, sinérgica, asociación que Koestler denomina una **bisociación**. (Koestler, p 35). El acto de creación, en este caso, la creación de una situación de humor, siempre opera en más de un marco de referencia. Las rutinas de pensamiento en un mismo marco de referencia se mencionan como de sentido único; en cambio, las que resultan de la intersección de dos marcos de referencia incompatibles, se denominan de sentido doble. Según Koestler, estas últimas actividades mentales son estados transitorios de equilibrio inestable donde el balance de las emociones y pensamientos está perturbado y que podrían jugar un rol determinante en la creación, tanto en ciencias como en arte; sin duda, en el humor [6].*

A continuación, una traducción libre del texto de Koestler, (1964 pp. 105-108), en el cual, el autor interpreta el famoso descubrimiento de Arquímedes, a través del marco conceptual de las matrices de pensamiento y que podría arrojar luces acerca de cómo toma lugar la interpretación en contexto de los resultados de datos procesados a través de un procedimiento estadístico [3].

*A Hierón, tirano de Siracusa y protector de Arquímedes, le había sido dada una hermosa corona, supuestamente de oro puro, pero que, el sospechaba, estaba adulterada con plata. Preguntó su opinión a Arquímedes. Arquímedes conocía, por supuesto, el peso específico del oro, esto quiere decir, su peso por unidad de volumen. De manera que, si él hubiese podido medir el volumen de la corona, él sabría inmediatamente si era oro puro o no; pero ¿cómo alguien podría determinar el volumen de un complicado adorno con toda su filigrana?*

*Podemos imaginar los pensamientos de Arquímedes moviéndose en círculos dentro del marco de su conocimiento geométrico; encontrando todas las aproximaciones al elusivo objetivo, el cual permanecía inalcanzable, y regresando una y otra vez al punto de partida. Esta frustrante situación, familiar para cualquiera que intente resolver un problema difícil, podría esquematizarse como en el diagrama de Figura 2, donde S representa el punto de partida, los bucles son trenes de pensamiento dentro de la matriz desde la cual no hay acceso aparente a la solución, y T representa el objetivo (esto es: ‘un método para medir el volumen de la corona’), que, desafortunadamente, se encuentra fuera del plano de la matriz [3].*

*Un día, mientras se bañaba, Arquímedes observó distraídamente la familiar visión del nivel del agua subiendo en la bañera como resultado de la inmersión de su cuerpo y, en un instante, se le ocurrió que el volumen de agua desplazado era igual al volumen de las partes sumergidas de su propio cuerpo, el cual, por tanto, podía medirse en forma simple.*

*Intentemos ponernos en el lugar de Arquímedes. Él tenía la costumbre de tomar habitualmente un baño, pero las expe-*

riencias e ideas asociadas con este acto se movían a lo largo de caminos trillados por el hábito. Ni a Arquímedes ni a nadie antes que él se le había ocurrido alguna vez conectar la ocupación sensual y trivial de tomar un baño caliente con la búsqueda académica de la medición de sólidos. Sin duda había observado muchas veces que el nivel del agua subía cada vez que entraba en la bañera; pero este hecho, y la distancia entre los dos niveles, eran totalmente irrelevantes para él —hasta que repentinamente este estado de cosas se **bisocia** con su problema. En ese instante se da cuenta de que la cantidad de aumento del nivel del agua era una simple medida del volumen de la parte sumergida de su cuerpo. El descubrimiento ahora puede esquematizarse como se presenta en Figura 3, donde M1 es la misma matriz que en el esquema anterior (ver Figura 2), regido por las reglas habituales del juego, por medio de las cuales Arquímedes intentó originalmente resolver el problema; M2 es la matriz de asociaciones relacionadas con el baño; m2 representa el tren de pensamientos que produce la conexión. El Enlace L pudo haber sido un concepto verbal (ver Koestler, p. 106); también pudo haber sido una impresión visual en la que de repente vio que el nivel del agua correspondía al volumen de las partes sumergidas del cuerpo y, por ende, a la de la corona, cuya imagen estaba constantemente acechando en el margen de su conciencia. El punto esencial es que en el momento crítico ambas matrices M1 y M2 estaban simultáneamente activas en la mente de Arquímedes, aunque presumiblemente en diferentes niveles de conciencia. El estrés creativo resultante de la situación no resuelta había mantenido el problema en la agenda aun cuando el rayo de la conciencia estaba a la deriva en otro plano completamente diferente. Sin esta presión constante, la constelación de posibilidades favorables (al descubrimiento) habría pasado desapercibida y se habría unido a la legión de oportunidades desaprovechadas por el hombre; oportunidades para un alejamiento creativo de los rancios hábitos de pensamiento que nublan sus poderes mentales [3].

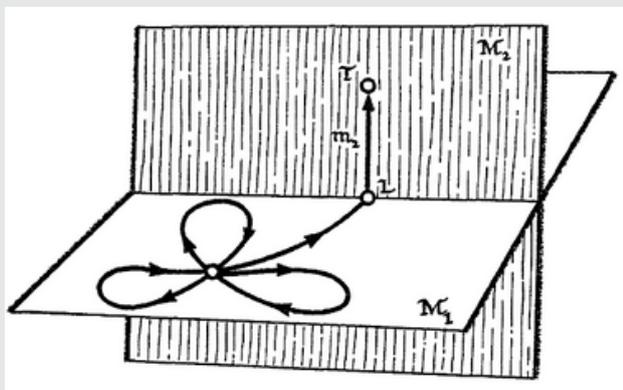


Figura 1.

Consideremos el descubrimiento de Arquímedes desde un ángulo diferente. Cuando uno se mete en una bañera uno sabe que el nivel del agua se elevará debido a su desplazamiento por el cuerpo, y que debe haber tanta agua desplazada como hay cuerpo sumergido; aún más, uno estima mecánicamente la cantidad de agua que se dejará entrar en la bañera debido a esta expectativa. También Arquímedes debía saber todo esto, pero probablemente nunca antes lo había verbalizado, es decir, nunca antes había formulado conscientemente ese fragmento de conocimiento. Sin embargo, implícitamente estaba allí como parte de su equipamiento mental; estaba, por así decirlo, incluido en el código de reglas para bañarse; en M2. Podemos ver que las reglas que gobiernan la matriz de una habilidad funcionan en un nivel de conciencia más bajo que el desempeño real en sí mismo, sea este tocar el piano, mantener una conversación o darse un baño. También podemos ver que el shock bisociativo a menudo tiene el efecto de hacer explícitas esas reglas implícitas, de repentinamente focalizar la conciencia en aspectos de la experiencia que no habían sido verbalizados,

inconscientemente implícitos, dados por sentado; de modo que un aspecto familiar e inadvertido de un fenómeno —como el aumento del nivel del agua— se percibe repentinamente desde un ángulo desconocido y significativo. 'Descubrimiento' a menudo significa simplemente descubrir algo que siempre ha estado ahí pero que estaba oculto a los ojos por las anteojeras de la costumbre. Esto se aplica igualmente a los descubrimientos del artista que nos hace ver objetos y acontecimientos familiares bajo una luz extraña, nueva y reveladora, como si atravesara la catarata que oscurece nuestra visión [3].

Hasta aquí, la traducción libre. De manera que la interpretación en contexto de datos ya procesados a través de un procedimiento estadístico, va en la búsqueda de estas bisociaciones, las que requieren al menos dos matrices de pensamiento activas simultáneamente en la mente del investigador.

Koestler señala, en páginas 39, 42, 44 y 45:

(...) en las rutinas del pensamiento disciplinado, sólo una matriz está activa en un momento. (...) Habilidades mentales tales como operaciones aritméticas, habilidades motoras tales como tocar piano o la escritura táctil tienden a ser, con la práctica, más o menos automatizadas, convirtiéndose en rutinas incorporadas que se gatillan por medio de 'señales codificadas' a través del sistema nervioso.

(...) Las matrices (de pensamiento) varían en flexibilidad, desde reflejos y rutinas más o menos automatizadas que permiten la elección de algunas posibles estrategias, hasta habilidades de una variedad casi ilimitada; pero, todo pensamiento o comportamiento coherente está sujeto a algún código de reglas posible de ser especificado y que es el responsable de su coherencia; aun cuando este código funcione parcialmente o enteramente, en niveles inconscientes de la mente, como generalmente ocurre.

(...) Por otro lado, el pensamiento que permanece confinado en una sola matriz, tiene limitaciones obvias. (...) Hay dos maneras de escapar a nuestras más o menos automatizadas rutinas de pensamiento y comportamiento. La primera es sumergirse en el sueño o en estados de ensoñación, cuando los códigos de pensamiento racional son suspendidos. La otra manera es también un escape (...) pero un escape en la dirección contraria; son las señaladas por flashes espontáneos de visión (insights) que muestran una situación familiar o evento bajo una nueva luz, y que inspiran una nueva respuesta. El acto bisociativo conecta matrices de experiencia previamente no conectadas.

(...) Cuando dos matrices independientes de percepción o razonamiento interactúan una con la otra, el resultado es una colisión que termina en una risa, o en su fusión que es una nueva síntesis intelectual, o en su confrontación en una experiencia estética [3].

J. Tal profundiza en las condiciones necesarias para que la interpretación en contexto de los resultados de los procedimientos estadísticos tome lugar [5]. Retomemos la cita de Tal en el inicio:

La descripción estadística involucra caracterizar los datos. No buscamos entender la razón por la que aparecen como aparecen. Por el contrario, la explicación estadística ocurre cuando, para usar las palabras de Koestler, 'dos matrices diferentes' de fenómenos se conectan. Es salir de un fenómeno particular para comprenderlo. Aparentemente existe una inclinación natural a ir más allá. (...) En el ámbito estadístico, este 'ir más allá' implica pasar de la mera descripción a la explicación. Esta relación de diferentes contextos entre sí es la esencia de la estadística inferencial. Como el término lo implica, hacemos uso de esta disciplina para inferir—comprender— por qué un conjunto de datos

J. Tal profundiza en las condiciones necesarias para que la interpretación en contexto de los resultados de los procedimientos estadísticos tome lugar [5]. Retomemos la cita de Tal en el inicio:

*tiene el aspecto que tiene. Más específicamente, al hacer estadística inferencial, nos preguntamos por qué nuestros fenómenos no son todos iguales. (...) La estadística inferencial es la disciplina de explicar la variación. [5]*

De manera que la interpretación en Estadística requiere ir más allá de los datos, en búsqueda de **bisociaciones que ocurran en la mente del investigador entre las matrices de pensamiento estadístico y las matrices de pensamiento a las que pertenecen las unidades de análisis** que dieron origen a los datos.

## EL ROL DE LA INTUICIÓN

Continúa J. Tal, (pp. 125-126):

*(...)Sin embargo, las estadísticas inferenciales no pueden seleccionar conexiones. Esto debemos hacerlo nosotros mismos. Evidentemente, la inclinación a 'ir más allá' no es suficiente. Si deseamos llegar, será mejor que primero sepamos adónde queremos ir; debemos, con previsión, seleccionar las variables que mostrarán el camino.*

*Pero la recopilación de datos es una actividad discriminatoria, como recoger flores y, a diferencia de la acción de una cortadora de césped, la selección de las flores que se consideran dignas de ser recogidas, así como su disposición en un ramo, son en última instancia cuestiones de gusto personal.*

*Así que incluso aquí el 'gusto personal', el pensamiento de 'sopa primordial', desempeña un papel. **Es la intuición y no las estadísticas la que determinará si vamos a hacer un descubrimiento sustancial o simplemente demostrar una conexión obvia.***

*(...) Al hacer estadísticas utilizamos el pensamiento analítico: recopilamos datos y especificamos modelos. Usamos la lógica. No puede ser de otra manera. Al mismo tiempo, si renunciamos a nuestra intuición y nos quedamos únicamente con nuestros modelos estándares, estaremos destinados para siempre a descubrir sólo conexiones triviales.*

*Un científico común y corriente no espera hacer descubrimientos trascendentales a diario. De hecho, si hace uno durante su vida, se considera afortunado. Por lo tanto, en lo principal, sigue las reglas. Sin embargo, si sólo hace esto, nunca irá mucho más allá de lo común. [5]*

Cabe destacar que "primordial soup" es el término coloquial que utiliza Tal para referirse a la intuición. Por tanto, el rol de la intuición es primordial [5]. Es la intuición la que conduce la selección de las variables que serán fruto de las mediciones. Koestler, p. 233, profundiza en este aspecto citando, entre otros, a T. H. Huxley:

*Como T. H. Huxley ha dicho en un pasaje frecuentemente citado:*

*Aquellos que se niegan a ir más allá de los hechos rara vez llegan más lejos que los mismos hechos; y cualquiera que ha estudiado la historia de la ciencia sabe que casi cada paso en ella ha sido dado por (...). la invención de una hipótesis que, aunque verificable, a menudo tenía poco fundamento para comenzar con ella (...). [3].*

Continúa Koestler, en la página 234:

*Constantemente surgen nuevos hechos; pero ellos se encuentran como resultado de una búsqueda en una direcc-*

*ción definida, basada en consideraciones teóricas –como Galle descubrió el planeta Neptuno, el cual nadie había visto antes, dirigiendo su telescopio a la región celeste que se trata de un caso extremo de observación guiado por la teoría; pero sigue siendo cierto que **no basta con que el científico mantenga los ojos abiertos a menos que tenga una idea de lo que él está buscando** [3].*

A continuación, dos ejemplos presentados por Koestler, p. 234, que nos muestran que no basta con los datos:

*En lo que respecta al conocimiento real, Copérnico no estaba en mejor situación (...) que los astrónomos griegos de Alejandría que vivieron en la época de Jesucristo. Tenían los mismos datos, los mismos instrumentos, los mismos conocimientos en geometría, al igual que él. Eran gigantes de la 'ciencia exacta'; sin embargo, no lograron ver lo que Copérnico vio después y Aristarco había visto antes que ellos: que los movimientos de los planetas estaban obviamente gobernados por el sol.*

*(...) Einstein formuló su 'Teoría Especial de la Relatividad' en 1905 basándose en datos que, como ya he dicho, no eran en absoluto nuevos. Poincaré, por ejemplo, veinticinco años mayor que Einstein, había tenido todos los hilos sueltos en sus manos y las razones por las que no logró unirlos siguen siendo materia de especulación entre científicos [3].*

Hevia, 2023, p. 185, señala:

*El significado estadístico que se obtiene de la aplicación de un procedimiento estadístico, pertenece al universo de discurso de los estadísticos. Los paradigmas del pensamiento estadístico toman lugar en ese orden de cosas y se expresan en lenguaje especializado; en lenguaje de expertos [6].*

Por tanto, **es tarea inexcusable del estadístico realizar la interpretación en contexto**; pero, también es de su responsabilidad la definición acertada de las variables que permitirán dejar en evidencia el fenómeno que se busca revelar a través de una elaboración del pensamiento estadístico, de un artefacto estadístico. Este punto es decisivo ya que sus decisiones permitirán o no permitirán que el pensamiento estadístico demuestre su fuerza y contundencia. El investigador debe avanzar guiado por su visión, si pretende alcanzar la claridad que le permita diseñar la estrategia que conduzca a la consecución del éxito. Resulta entonces, que es **la visión la que, en definitiva, conducirá toda la implementación de la estrategia que llevará al investigador al éxito**. Es relevante, entonces, recorrer estos caminos a lo largo de los cuales los estadísticos, guiados por su visión, elaboran sus artefactos, se guarecen de los peligros y manteniendo a raya a sus principales enemigos avanzan de la mano de sus aliados, esos paradigmas que los sostienen en sus heroicas hazañas. Es lo que haremos a continuación.

## PRESENCIA DEL PENSAMIENTO ESTADÍSTICO

J. Tal en su texto *Reading Between the Numbers*, p. 14, utiliza la famosa investigación realizada por Sir Francis Galton sobre las estaturas de los seres humanos como factor de herencia, para ejemplificar la actividad que desarrolla un estadístico en sus investigaciones [5]. En particular, J. Tal identifica seis etapas en el desarrollo de una investigación de este tipo [5]:

1. Realice una pregunta
2. Diseñe un estudio y especifique sus medidas
3. Recopile datos y descríbalos
4. Simplifique los datos
5. Interprete los datos
6. Generalice los hallazgos

A continuación, revisamos algunos de estos aspectos estudiando el artículo *Regression towards Mediocrity in Hereditary Stature*

de Sir Francis Galton publicado en "The Journal of the Anthropological Institute of Great Britain and Ireland", 1886, pp. 246-263. [7]

### 1. Plantee un interrogante en términos de variables que recogen el efecto de interés

Galton estaba fuertemente interesado en las leyes de herencia, tal como lo declara expresamente en su artículo de 1886, donde también proporciona evidencia acerca de sus investigaciones eliminando sobre la reproducción de semillas:

*Mi objetivo es establecer más allá de toda duda la existencia de una ley simple y de largo alcance que regula la transmisión hereditaria de las cualidades que todos poseemos, aunque en grados desiguales. Alguna vez antes me aventuré a llamar la atención acerca de esta ley con una evidencia lejos más exigua que la que ahora poseo.*

*Hace ya algunos años que yo hice una extensa serie de experimentos sobre la producción de semillas de diferentes tamaños, pero de la misma especie. Ellos dieron resultados que parecían muy notables, y que usé como base para una conferencia ante la Royal Institution, el 9 de febrero de 1877. Aparecía en estos experimentos que la progenie no tendía a asemejarse a sus semillas parentales en tamaño, sino que siempre eran más mediocres que ellas –más pequeñas que sus padres si sus padres eran grandes; más grandes que los padres, si los padres eran muy pequeños.*

*(...) Los experimentos mostraron aún más que el promedio de la regresión filial hacia la mediocridad era directamente proporcional a la desviación parental correspondiente. Este curioso resultado estaba basado en muchas plantaciones realizadas para mí, por amigos que vivían en varias partes del país, ... por lo que no pude dar lugar a duda alguna a la verdad de mis conclusiones. La exacta razón de la regresión, permaneció un poco dudosa, debido a las influencias de variables; por tanto, no intenté definirla [7].*

La cuestión investigativa había estado en la mente de Galton por años [7]. J. Tal, p. 12, señala que Galton, en su libro *Hereditary Genius*, 1869, expresa:

*(...) comencé pensando en las actitudes y logros de mis contemporáneos en la escuela, en la universidad y en sus vidas posteriores, y me sorprendió encontrar cuan frecuentemente la capacidad parecía ir con la descendencia.*

*A lo largo de los años, recopiló datos sobre la inteligencia humana y concluyó que (...) los resultados fueron tales, en mi propia opinión, como para establecer completamente la teoría de que la genialidad era hereditaria' [5].*

### 2. Diseñe un estudio; especifique sus mediciones, cuide la posible presencia de sesgo en los datos, asegure la presencia de variabilidad significativa en los datos en el sentido de que esta variabilidad dé cuenta del efecto que se quiere evidenciar

Conforme narra Galton, p. 247 :

*Después de que se publicó la conferencia, se me ocurrió que los motivos de mis dudas podrían invocarse como objeciones a las conclusiones generales. No pensé en ellos en el momento, pero como la investigación se había visto rodeada de muchas pequeñas dificultades y cuestiones de detalle, difícilmente sería posible dar una breve y al mismo tiempo adecuada y completa respuesta a tales objeciones. ... así que pensé que era mejor no decir nada sobre el tema hasta que obtuviese evidencia independiente. Era evidencia antropológica la que yo deseaba, preocupándome por las semillas solo como medio para arrojar luz sobre la herencia en el hombre. Traté en vano por un largo y agotador tiempo de obtenerla en abundancia suficiente y haber fallado fue un*

*motivo poderoso, junto a otros, en inducirme a hacer una oferta de premios por Registros de Familia, la cual tuvo una gran respuesta y que me proveyó el último año con lo que yo quería. [7]*

Galton nos hace presente las dificultades a las que nos lleva la obtención de datos confiables, los que también deben ser apropiados para el estudio [7]. La genialidad de Galton resuelve el problema de la generación de los datos a través de estos Registros de Familia que se pedían completar y remitir a Galton, con el compromiso de ser retornados a cada familia que participaba en el estudio [7][8]. Observe que las unidades de análisis son los hogares del país; el marco muestral está constituido por aquellas familias que respondieron los registros de la familia, entregando información sobre la familia constituida y con el interés de ganar los premios ofrecidos. De esta forma, Galton obtiene respuestas completas de 205 hogares los que comprenden un total de 930 hijos o hijas adultos [7].

*Me cuidé especialmente de no hacer alusión alguna en mi prospecto a la investigación específica, para evitar la formación de sesgo en los registros devueltos. Ahora puedo contemplar con tranquilidad la posibilidad de que los registros de altura hayan sido elaborados frecuentemente de una manera descuidada, porque ninguna cantidad de inexactitud insesgada (error no aleatorio) puede dar cuenta de los resultados. Galton, p. 247. [7].*

En relación al error que contienen los datos, Galton, (p. 252), nos entrega la siguiente cota superior para el error en las mediciones de las alturas. Ello muestra que existe variabilidad significativa por sobre la variabilidad atribuible al error aleatorio:

*La única desventaja en el uso de la estatura es su pequeña variabilidad. La mitad de la población con la que traté variaba menos de 1,7 pulgadas del promedio de todos ellos (...) Por otro lado, (...) parcialmente debido a la incertidumbre (...) de si las alturas fueron medidas con los zapatos puestos o sin zapatos, yo encontré por medio de una investigación independiente (...) que cada observación está sujeta a un error que no excede 2/3 de una pulgada [7].*

Finalmente, el siguiente párrafo justifica bellamente la elección de la estatura como variable clave que da cuenta de "la existencia de una ley simple y de largo alcance que regula la transmisión hereditaria de las cualidades que todos poseemos, aunque en grados desiguales". Ver Galton, p. 247:

*Ahora explicaré con detalle por qué elegí la estatura como sujeto de investigación; haciendo esto se manifestarán mejor las peculiaridades y puntos a ser atendidos en la investigación. Muchas de sus ventajas son bastante obvias, como la facilidad y frecuencia con la que se realiza su medición, su constante práctica durante treinta y cinco años de una vida promedio, su poca dependencia de las diferencias en la educación, y su no importante influencia en la tasa de mortalidad. Otras ventajas que no son tan obvias no son menos grandiosas. Una de ellas se encuentra en el hecho de que la estatura no es un simple elemento, sino una suma de longitudes acumuladas o espesores determinados de más de cien partes del cuerpo, cada una tan distinta del resto como para haberse ganado un nombre por el cual puede ser especificada. La lista de ellas incluye alrededor de cincuenta huesos separados, situados en el cráneo, la columna, la pelvis, las dos piernas y los dos tobillos y pies. Se cuentan los huesos de ambas extremidades inferiores, porque es la longitud promedio de estas dos extremidades la que contribuye a la estatura general. Los cartílagos interpuestos entre los huesos, dos en cada articulación, son bastante más numerosos que los propios huesos. Las partes carnosas del cuero cabelludo de la cabeza y de las plantas de los pies concluyen la lista. Se debería tomar en cuenta también la forma y ubicación de muchos de los huesos que conducen a un empeine más o menos arqueado, espalda*

recta o cabeza alta [7].

### 3. Recopile datos y descríbalos

Dejemos que Galton, p. 247, describa sus datos. Observe cómo Galton aprovecha los resultados de una indicación errada en los cálculos, para dar evidencia de la robustez de sus procedimientos:

*Mis datos consistieron en las estaturas de 930 niños adultos y de sus respectivos padres, 205 en número. En cada caso transmuté las estaturas femeninas a sus correspondientes masculinos equivalentes y las usé en su forma transmutada, de modo que no hubiese objeción basada en la diferencia sexual de estatura que necesitase ser considerada cuando hablo de promedios. El factor que utilicé fue 1,08, que equivale a añadir un poco menos de una doceava parte a la altura de cada mujer. (Este factor) difiere muy poco de los factores empleados por otros antropólogos, quienes, aún más, difieren un poco entre sí; de todos modos, se adapta mejor a mis datos que 1,07 o 1,09. El resultado final no es del tipo que se vea afectado por estos minuciosos detalles, porque sucedió que, debido a una indicación equivocada, el calculista a quien le confié por primera vez las cifras usó un factor algo diferente, sin embargo, el resultado fue casi el mismo [7].*

A continuación, Galton observa un hecho que resulta ser de gran conveniencia para su estudio: la estatura de los hijos depende muy cercanamente del **promedio** de estaturas de sus dos padres pudiendo ser considerada, en la práctica, como no teniendo relación alguna con las estaturas de sus padres por separado [7]. Para demostrar este hecho, Galton ordena en forma acumulada y creciente las alturas de los hijos adultos que difieren 1, 2, 3, 4, etc. pulgadas del promedio de altura de los hijos de su respectiva familia, agrupándolos en segmentos según sean las diferencias de altura entre sus padres: Menos de 1 pulgada; 1 pulgada y menos que 2 pulgadas; 2 pulgadas y menos que 3 pulgadas; etc [7]. Luego, compara estas cifras en una tabla cuyas entradas corresponden al número respectivo de casos sobre un total de 50 casos. Para cuidar la confiabilidad del promedio de altura de los hijos de cada familia, Galton considera sólo familias que tengan 6 hijos o más. El gráfico de la Figura 2 muestra la regularidad de la distribución de las estaturas, según los segmentos definidos por diferencias de alturas entre los padres. De hecho, Galton prueba que la distribución de las diferencias de estaturas de los hijos adultos respecto a la media familiar de los hijos, es independiente de las diferencias entre las estaturas de sus padres. (Actualmente, nosotros aplicaríamos una prueba Chi cuadrado de homogeneidad para el caso.) No deja de ser interesante el comentario que Galton agrega acerca de su Tabla II en relación al último segmento:

*Las entradas en cada una de las diferentes líneas se ven comportándose de la misma manera, excepto que en la última de ellas los niños mostraron una tendencia débil a dividirse en dos grupos, uno siguiendo a los padres altos, el otro siguiendo a los más bajos. Esto, sin embargo, no es visible en el resumen de Tabla II que anexo. Por lo tanto, cuando se trata de la transmisión de estatura de padres a hijos, el promedio de altura de los dos padres, o, como prefiero llamarlo, la 'altura media parental', es todo lo que necesitamos saber sobre ellos [7].*

Finalmente, una última descripción de los datos, permite a Galton asegurar que las estaturas por género no están asociadas a la formación de los matrimonios de las familias [7].

*(...) mi test consistió en dividir a los 205 padres varones y las 205 madres, cada uno en tres grupos: T, M y S, es decir, alto (tall), mediano (medium) y bajo (short) (las medidas masculinas medianas se toman desde 67 pulgadas hacia arriba hasta 70 pulgadas), y en contar el número de matrimonios en cada combinación posible entre ellos. El resultado fue que hombres y mujeres de alturas de contraste, baja y alta o alta y baja, se casan casi tan frecuentemente*

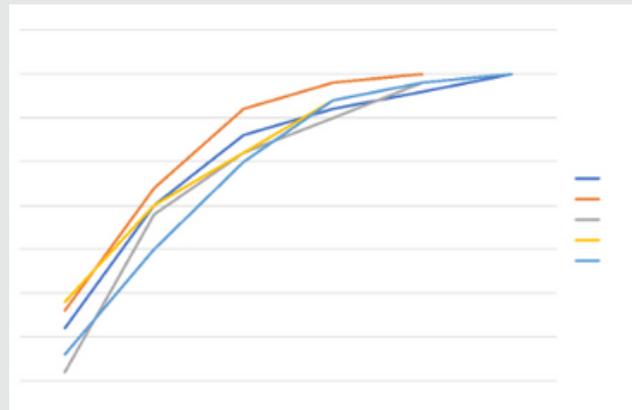


Figura 2.

*como hombres y mujeres de estatura similar, ambas altas o ambas bajas; Hubo 32 casos de uno y 27 casos del otro [7].*

Cabe destacar que habría un error de cálculo en la cifra "32", siendo 26 la correcta.

Aplicando el test no paramétrico de independencia basado en el estadístico Chi cuadrado, se obtiene un valor p igual a 0,5735; confirmándose esta última observación de Galton.

### 4. Simplifique los datos

Veamos este paso siguiendo la narración de Galton, p. 254.

*(...) (Tabla I) se deduce de una gran hoja en la que anoté la altura de cada hijo, en oposición a su altura media parental, y en cada caso, cada uno fue ingresado aproximándolo a la décima de pulgada más cercana. Luego, conté el número de entradas en cada pulgada cuadrada y los transcribí tal como aparecen en la tabla. Se entiende mejor el significado de la tabla por medio de ejemplos. Así, de un total de 928 hijos que nacieron de los 205 padres promedio en mi lista, había 18 con una altura de 69,2 pulgadas (contando hasta la pulgada más cercana), quienes hubieron nacido de padres promedio de una altura de 70,5 pulgadas (también contando hasta la pulgada más cercana). Así nuevamente, hubo 25 niños de 70,2 pulgadas nacidos de padres promedio de 69,5 pulgadas. Al principio me resultó difícil captar todo el significado de las entradas de la tabla, la cual tenía relaciones curiosas que fueron muy interesantes de investigar. Ellas se mostraron claramente cuando 'suavicé' las entradas escribiendo en cada intersección de una columna horizontal con una vertical, la suma de las entradas en los cuatro cuadrados adyacentes, y usarlas para trabajar en ellas [7].*

	Below	62,2	63,2	64,2	65,2	66,2	67,2	68,2	69,2	70,2	71,2	72,2	73,2	Above
Above	0	0	0	0	0	0	0	1	2	1	3	10	3	11
72,5	0	0	0	0	1	3	5	5	7	14	12	14	16	23
71,5	1	0	1	1	4	6	9	22	33	24	28	17	16	52
70,5	0	2	1	18	6	24	44	44	64	60	42	30	13	78
69,5	2	1	24	16	50	57	71	106	111	88	61	32	22	190
68,5	0	11	17	53	55	100	124	127	126	110	56	36	8	227
67,5	4	8	27	36	68	95	112	124	108	74	41	15	7	211
66,5	4	6	22	24	44	66	80	65	63	37	17	5	0	78
65,5	1	14	12	25	19	40	40	32	27	13	6	2	1	66
64,5	3	5	16	14	17	19	18	15	8	7	2	1	0	23
Below	1	4	8	7	7	8	8	3	3	1	0	0	0	14

Figura 3.

Galton continúa sus comentarios sobre aquello que visualiza en la tabla suavizada después de haber trabajado en sus entradas:

*Entonces me di cuenta que las líneas dibujadas a través de entradas del mismo valor formaban una serie de elipses concéntricas y similares. Su centro común estaba en la intersección de las líneas verticales y horizontales, que correspondían a 68 1/4 pulgadas. Sus ejes estaban similarmente inclinados. Los puntos donde cada elipse en sucesión era tocada por una tangente horizontal, yacían en una línea recta inclinada a la vertical en la razón 2/3; aquellos tocados por una tangente vertical yacían en una línea recta inclinada a la horizontal en la razón 1/3. Estas relaciones confirman los valores de la regresión promedio ya obtenidos por un método diferente, de 2/3 desde los padres promedio hasta la descendencia, y de 1/3 desde la descendencia a los padres promedio (...) [7].*

**5. Interprete los datos**

Efectivamente, utilizando los valores marginales de Tabla I de Galton, se pueden obtener las relaciones entre “desviaciones” que establece su ley, tal como Galton señala en página 252:

*La ley que deseo establecer se refiere primariamente a la herencia de diferentes grados de ser alto y ser bajo, y sólo en forma secundaria a la estatura absoluta. Es decir, se refiere a medidas realizadas desde la corona de la cabeza hasta el nivel de la media (mediocridad), hacia arriba o hacia abajo según sea el caso, y no desde la corona de la cabeza hasta el suelo. En la población con la que estoy tratando, el nivel medio es de 6814 pulgadas (sin zapatos). La misma ley aplicada con suficiente cercanía, tanto en altas estaturas como en bajas estaturas, podría incluir ambas bajo el único encabezado de desviaciones, y llamaré ‘desviación (deviate)’ a cualquier desviación particular. A través del uso de esta palabra y la de ‘padres promedio’ podemos definir la ley de regresión muy brevemente. Es que la desviación de la descendencia es, en promedio, dos tercios de la desviación de sus padres promedio.*

*Lámina IX, fig. a, da una expresión gráfica de los datos sobre la que se fundamenta la presente ley. En ella se verá que las relaciones entre la estatura de los hijos y la de sus padres promedio, que son perfectamente simples cuando se refieren a la escala de desviaciones a la derecha de la platina, no admiten ser redactados brevemente cuando se refieren a la escala de estaturas a su izquierda [7].*

Los “pequeños puntos” graficados tienen: coordenadas y desviaciones respecto a 68 1/4 como se indica en la tabla de

Figura 4. El promedio de las desviaciones (excluyendo el centro) es 0,65; es decir, aproximadamente 2/3 desde el padre medio hasta la descendencia.

Galton, en página 247, la segunda página de su artículo, ya había anunciado estos resultados:

*El análisis de los Registros (de Familia) confirma plenamente y va mucho más allá de las conclusiones que obtuve de las semillas. Entrega el valor numérico de la regresión hacia la mediocridad en el caso de la estatura humana, desde 1 hasta 2/3 con coherencia y precisión inesperadas [ver Lámina IX, fig. (a)], y me proporcionó la clase de hechos que yo quería investigar –los grados de parecido familiar en diferentes grados de parentesco y los pasos a través de los cuales las peculiaridades especiales de una familia se fusionan finalmente con las características típicas de la raza [7].*

En la página 261, Galton proporciona varias explicaciones acerca de la Lámina IX, Fig. a; y que permiten clarificar mejor sus procedimientos.

- *Las líneas horizontales cortas se refieren a la estatura de los padres promedio como se indica en la escala de la izquierda. Estos son los mismos valores que los de la columna izquierda del Cuadro I. Los círculos pequeños, uno debajo de cada uno de los anteriores (valores), muestran la estatura media de los hijos de cada uno de esos padres promedio. Estos son los valores en la columna de la derecha de la Tabla I, titulada ‘Medianas’. (La Mediana es el valor que excede a la mitad de los casos, y que no alcanza a la otra mitad. Es prácticamente lo mismo que la media, pero es un valor más conveniente de utilizar, en la forma de trabajo adoptada a todo lo largo de la presente instancia) [7].*

**6. Generalice los hallazgos**

Aparentemente, aún se discute si los resultados de Galton dan evidencia de “la existencia de una ley simple y de largo alcance que regula la transmisión hereditaria de las cualidades que todos poseemos, aunque en grados desiguales” [6]. Sin embargo, sus estudios no se detienen en estos resultados, los que hemos analizado con el propósito de dejar en evidencia aquellos aspectos del pensamiento que dan cuenta de cómo se trabaja en la ciencia de los datos. Las generalizaciones que se extienden de esta investigación, también tienen umbilicaciones en aspectos de la Estadística Matemática de amplia discusión en ese momento y que se dejan entrever en sus estudios de las secciones de cierta superficie de frecuencia de error, que darían cuenta de las elipses tobobservadas en sus datos (ver Galton, p. 255); sin extenderse en su acertada conjetura sobre la forma matemática del coeficiente de regresión, aparentemente, desconociendo el método de mínimos cuadrados ya desarrollado por Gauss algunos años antes [7][9].

Pequeños puntos										
abscisa	64,5	65,5	66,5	67,5	68,5	69,5	70,5	71,5	72,5	
ordenada	65,8	66,7	67,2	67,6	68,2	68,9	69,5	69,9	72,2	
Desviaciones										
abscisa	-3,75	-2,75	-1,75	-0,75	0,25	1,25	2,25	3,25	4,25	
ordenada	-2,45	-1,55	-1,05	-0,65	-0,05	0,65	1,25	1,65	3,95	
razón	0,65	0,56	0,6	0,87	-0,2	0,52	0,56	0,51	0,93	razón promedio
										0,65

Figura 4

## OBSERVACIONES FINALES

A lo largo de este trabajo, uno de mis objetivos ha sido establecer claramente la diferencia que existe entre “interpretación en contexto” y “pensamiento estadístico en sí”, como he planteado en Hevia, 2023. Sin embargo y paradójicamente, la interpretación de los resultados de un procedimiento estadístico debe ser realizada por un estadístico; **es el único que a entera cabalidad conoce de este tipo de procedimientos**. Por tanto, es muy posible que esta paradójica situación sea la que explique la existencia de estadísticos en todas las áreas del saber, por muy diversas que estas sean y que han otorgado a la Estadística su maravillosa riqueza y universalidad. Otro objetivo ha sido continuar la tarea de reconocimiento de los drivers que conducen el pensamiento estadístico en distintos niveles de funcionamiento cognitivo en los que opera; tarea iniciada en Hevia, 2022 y definida explícitamente en Hevia, 2023 [10] [6].

Finalmente, me ha parecido muy apropiado cerrar estas ideas presentando dos comentarios de D. S. Moore, 1998, pp. 1253, 1258; en *Statistics Among the Liberal Arts*, que ameritan una mayor reflexión:

*El fracaso de la estadística a la hora de lograr un amplio reconocimiento como campo independiente es el problema más grave que enfrenta nuestra disciplina. Gerry Hahn y Roger Hoerl (1998) describen el entorno actual en la industria como ‘estadísticas sin estadísticos’, y lo mismo ocurre en el mundo académico. Cuando ‘realizar un análisis estadístico ya no es una tarea comercializable; cualquiera con una computadora portátil puede hacerlo’, ¿qué tienen los estadísticos para ofrecer al investigador médico, al psicólogo o al ingeniero? Ésta es la forma práctica de la pregunta de si la estadística es de hecho una disciplina separada y fundamental. Hahn y Hoerl dan la respuesta en un eslogan que debería ser el lema de la instrucción inicial: ‘pasar de los métodos estadísticos al pensamiento estadístico’ [1].*

*Debido a que las estadísticas implican formas de pensar distintivas y poderosas, no seremos absorbidos por la tecnología de la información. La revolución informática y de información muy desordenada. El pensamiento estadístico ofrece herramientas mentales simples, pero no intuitivas para recortar esta masa, ordenar el desorden, separar el sentido del sinsentido, seleccionar lo poco que es relevante de lo mucho que es irrelevante. Es más, cualquier revolución implica que enfrentemos nuevos problemas. Los nuevos problemas requieren modos de razonamiento generales y flexibles. Como lo expresó Bob Hogg (1989), **la estadística es una guía hacia lo desconocido** [1].*

## REFERENCIAS BIBLIOGRÁFICAS

1. Bruner JS. The course of cognitive growth. *Am Psychol.* 1964;19(1):1-15. <https://doi.org/10.1037/h0049682>
2. Moore DS. Should mathematicians teach statistics? *Coll Math J.* 1988;19(1):3-7. <https://doi.org/10.1080/07468342.1988.11973704>
3. Koestler A. *The act of creation*. London: Hutchinson & Co.; 1964.
4. Langdridge D. *Phenomenological psychology: Theory, research and method*. Harlow: Pearson Prentice Hall; 2007.
5. Tal J. *Reading between the numbers: Statistical thinking in everyday life*. New York: McGraw-Hill; 2001.
6. Hevia H. Pensamiento Estadístico: ¿Qué aportan las Matrices de Koestler? In: Sagula JE, editor. *Memorias del IV Simposio de Educación Matemática Virtual, Paradigmas evolutivos en Educación Matemática: Conferencias y artículos*. Lujan: EdUnLu; 2023. p. 1-10. ISBN 978-987-3941-87-0.
7. Galton F. Regression towards mediocrity in hereditary stature. *J Anthropol Inst Great Brit Ireland.* 1886;15(4):246-63. <https://doi.org/10.2307/2841652>
8. Record of Family Faculties [Internet]. Available from: <https://www.medicine.mcgill.ca/epidemiology/hanley/galton/RF/index.html>
9. Read C. *The econometricians: Gauss, Galton, Pearson, Fisher, Hotelling, Cowles, Frisch, and Haavelmo*. London: Palgrave Macmillan; 2016.
10. Hevia H. La idea de modelo de probabilidad de una población. In: *Cutting-edge research in mathematics and its applications*. Atena Editora; 2022. <https://doi.org/10.22533/at.ed.575221502>. ISBN 978-65-5983-957-5.
11. Moore DS. Should mathematicians teach statistics? *Coll Math J* 1988;19(1):3-7. <https://doi.org/10.1080/07468342.1988.11973704>