

# LA PSICOLOGÍA Y EL ESTADÍSTICO INTUITIVO

Alejandro A. Lazarte

## *LAS METÁFORAS EN CIENCIA*

Las metáforas han cumplido un rol muy importante en el desarrollo de todas las ciencias. Por ejemplo, Charles Darwin comparó la selección natural como una «guerra» por la supervivencia. En el caso de la Psicología, muchas metáforas han sido utilizadas para representar procesos psicológicos; por ejemplo, Platón comparaba la impresión de una «memoria» en la «mente» con la impresión que hacía un sello o un estilo sobre una tablilla de cera. San Agustín asemeja la memoria como un almacén en donde «almacenamos» información permanentemente, y el olvido estaría dado no por la desaparición de la información del almacén, sino a la probabilidad de llegar a localizarla de nuevo dentro del almacén. Esta última es muy similar a la metáfora de la computadora, que dio origen a fines de los 50s a lo que se denominó después la «revolución cognitiva». Hoy en día se habla de la metáfora neuronal, donde los procesos psicológicos son modelados por redes de «neuronas» conceptuales que caracterizan su función a través de las conexiones que se activan entre ellas.

Para Gigerenzer & Murray (1987), sobre los cuales basaré gran parte de esta presentación, una metáfora consiste de un *sujeto* y un *modificador*. Por ejemplo, el hombre es una máquina neumática. Pero a diferencia de una frase literal, existe una relación entre sujeto y modificador que no es la esperada por la intuición común, y que invita a mayor elaboración: por ejemplo, decir «el ser humano es un mamífero» es más una descripción que una metáfora, mientras que decir «contestar pruebas psicométricas es una carrera de caballos» invita a una mayor aclaración de sus implicaciones. En contraposición con simples enunciados raros, una metáfora es un enunciado inteligible y aceptable, aunque algo extraño. Contestar pruebas es, usualmente, correr en contra del tiempo, compitiendo con otras personas o con uno mismo, por lo tanto algo de común tiene con la idea de una competencia de caballos.

Las metáforas en sí no son falsificables. No podemos decir que una metáfora es válida, o no, sólo enunciándola. Sin embargo, especificando las implicaciones y las connotaciones de la metáfora podemos convertirla en enunciados precisos que sean posibles de ser puestos a prueba empíricamente. En esta transformación la metáfora pasa a generar un *modelo*, el cual ha perdido ya la vaguedad de la metáfora, es mucho más elaborado conceptualmente, y ofrece predicciones que pueden ser comprobadas por experimentación y/o observación. Es obvio que la utilidad de la metáfora se halla en estimular la construcción de teorías, más que en indicar cómo se pueden probar esas teorías.

Las metáforas en Psicología han cambiado a través del tiempo influenciadas en parte por las metáforas utilizadas en ciencias consideradas como más «avanzadas», y en parte por las innovaciones tecnológicas continuas que han caracterizado el presente siglo, y que han producido herramientas útiles para el trabajo cotidiano. Por ejemplo, el desarrollo y popularización del telégrafo y el teléfono llevó a la comparación entre una persona y un sistema de comunicación (Attneave, 1959). La invención de la computadora influyó enormemente la Psicología. A través de ella, muchas metáforas cognitivas pudieron por primera vez ser implementadas. Por ejemplo, el planeamiento y ejecución de comportamientos se comparó a la ejecución de un programa de computadora, recordar un evento se comparó con la búsqueda de la información en una base de datos, etc. Newell & Simon (1972) examinaron exhaustivamente las posibilidades de imitar el proceso de solución de problemas a través de programas de computadoras.

Paralelamente a la computadora, otra gran influencia en la generación de metáforas en Psicología ha sido otra herramienta que fue adoptada por los psicólogos en la década de los 40: la Estadística Inferencial de Ronald A. Fisher (1925), y Jerzy Neyman y Egon S. Pearson (1928a, 1928b).

Entre 1940 y 1955 las teorías de estadística inferencial se volvieron herramientas indispensables para realizar inferencias de los datos a las hipótesis dentro de Psicología (Cowles, 1989). La influencia de la estadística inferencial, y su institucionalización como la lengua franca para la comunicación de resultados experimentales dentro de Psicología, ha producido una forma particular de plantear preguntas de investigación y de diseñar la operacionalización de estas preguntas con miras a la contrastación empírica de las mismas. Este enorme prestigio de la Estadística se muestra claramente en el entrenamiento de estudiantes de doctorado en Psicología y Educación, y en los criterios que los revisores de artículos de revistas especializadas utilizan para seleccionar aquellos trabajos que ameritan publicación.

Siguiendo a Gerd Gigerenzer (Gigerenzer, 1987; Gigerenzer & Murray, 1987; Gigerenzer, Switjink, Porter, Daton, Beatty & Kruger, 1989) presentaré cómo la Estadística inferencial llegó a tener tal preeminencia dentro de la Psicología, y cómo, a partir del prestigio adquirido como herramienta metodológica, llegó a convertirse en una metáfora que fue utilizada para desarrollar teorías sobre procesos cognitivos, es decir, como se llegó a producir la metáfora del ser humano como un «estadístico intuitivo».

### ***EMERGENCIA DE LA ESTADÍSTICA INFERENCIAL***

El problema de la inferencia reside en cómo obtener conclusiones válidas y generales a partir de un conjunto particular de datos. Obviamente la inferencia de tipo estadístico no es la única forma de inferencia, por ejemplo es notable el poco uso y necesidad que físicos y químicos tienen de los procedimientos estadísticos. Hay dos polos en lo que respecta a la naturaleza de la inferencia inductiva. Por un lado, podemos considerar la inferencia de los datos a la hipótesis como un proceso cognitivo informal, basado en el juicio de un experto, y por lo tanto dependiente del contenido del problema y de la experiencia del individuo. Este tipo de enfoque es mantenido por físicos y científicos naturales.

El otro polo es considerar la inferencia como un proceso que puede ser descrito por una única regla formal, la cual puede ser aplicada independiente del contenido específico que está siendo investigado. Ha sido la teoría de la probabilidad la que ha brindado el único candidato a ese intento de crear una inferencia inductiva formalizada. El problema de la inferencia inductiva formalizada, o inducción, es relativamente reciente (Hacking, 1975). Frente a este problema han habido diferentes formas de atacarlo. Revisaremos brevemente las tres principales teorías de la inferencia estadística: La teoría Bayesiana, el Modelo de la Hipótesis Nula de Fisher, y el Modelo de toma de decisiones estadísticas de Neyman & Pearson.

Irónicamente, el enfoque Bayesiano es el que en mayor o menor medida todos los científicos tienen cuando intentan *interpretar* los resultados de una prueba estadística, aunque esta interpretación no sea admisible dentro de los métodos estadísticos usuales, que están basados en la tradición Fisheriana. Básicamente, el enfoque Bayesiano trata de obtener una estimación de la probabilidad de que una hipótesis sea verdadera, dado el tipo de datos que se ha observado en el experimento, es decir  $P(H|D)$ . Una aplicación muy sencilla del teorema de Bayes nos muestra que esta probabilidad es función de la *probabilidad a priori* de la hipótesis,

$P(H)$  y del impacto que tienen los datos sobre esta probabilidad, expresado como la razón entre la probabilidad de los datos dado que la hipótesis es verdadera y la probabilidad *a priori* de los datos:  $P(D|H)/P(D)$ . El aspecto controversial del enfoque Bayesiano reside en la introducción de la probabilidad *a priori* de la hipótesis, ya que para estimarla es necesario hacer una evaluación subjetiva de ella, algo que no es aceptado totalmente por la mayoría de estadísticos.

Ya que no es posible enfrentar el problema de la probabilidad de la hipótesis dados los datos, los estadísticos frecuentistas, o no Bayesianos, reformularon el problema de la inferencia estadística como el problema de *la probabilidad de los datos, dado que la hipótesis se asume verdadera*, es decir  $P(D|H)$ .

En la tradición Fisheriana, la *hipótesis nula* cumple el rol central en la inferencia estadística. Una hipótesis nula es un enunciado sobre algún parámetro poblacional, usualmente un promedio, que el experimentador utiliza como hipótesis de trabajo para generar los estimadores de la distribución muestral del estadístico a utilizar en la prueba. Esta hipótesis de trabajo es usualmente escogida por su conveniencia, más que por un verdadero interés por comprobarla. Es más, usualmente se construye de tal forma que vaya en contra de las predicciones de la teoría que uno quiere probar, y por lo tanto el investigador está usualmente interesado en rechazarla. Fisher nunca prestó mucha atención a que sucedía después de rechazar una hipótesis nula, ya que dentro de su enfoque no existía el concepto formal de una hipótesis alternativa que estaba siendo confirmada cuando la nula era rechazada.

El razonamiento a seguir para decidir si rechazamos la hipótesis nula es el mismo que se enseña en un curso introductorio de Estadística: asumiendo que la hipótesis nula es verdadera, calculamos la probabilidad de observar el valor del estadístico de prueba (el valor de una  $Z$ ,  $t$  o  $F$  observada); si esta probabilidad, o *valor  $p$* , es pequeña, decimos que el supuesto de que la hipótesis nula es verdadera no es creíble. Pero, ¿cuán pequeña tiene que ser esta probabilidad para rechazar la hipótesis nula? A falta de una conceptualización de hipótesis alternativa, Fisher sólo pudo brindar un criterio arbitrario para decidir cuando rechazar una hipótesis. Este criterio, el famoso *nivel de significancia o nivel alfa*, fue arbitrariamente fijado en .05 o en .01.

El enfoque de Neyman & Pearson, por el otro lado, es lo que se ha llegado a considerar como el modelo tradicional de teoría de decisión estadística. En este proceso de decisión hay *dos* hipótesis específicas, una nula y otra alternativa. Por lo tanto, ahora es factible definir dos posibles errores a cometer en cualquier decisión: (1) rechazar la hipótesis nula cuando ésta es verdadera o error tipo I,

cuya probabilidad es el alfa Fisheriana, y (2) aceptar la hipótesis nula cuando esta es falsa o error tipo II, cuya probabilidad se denomina beta. El complemento de beta, es decir  $1 - \beta$ , es el *poder estadístico* de la prueba. En este enfoque, para rechazar la hipótesis nula se debe tomar en consideración las probabilidades de los dos posibles tipos de errores, y *además* los costos que implica cada una de las decisiones. En este sentido, el problema matemático de evaluar las probabilidades de las hipótesis dados los datos, es obviamente distinto al proceso de decisión de rechazar o aceptar la hipótesis Nula. Neyman & Pearson consideran que estrictamente hablando, la hipótesis nula no es rechazada, sino que, basados en los resultados, *actuamos* como si la hipótesis alternativa es verdadera.

### ***LA INFERENCIA ESTADÍSTICA EN PSICOLOGÍA***

En la mayoría de la investigación psicológica contemporánea, los métodos de análisis de resultados están basados en el Análisis de Varianza de Fisher, donde la prueba de la hipótesis de igualdad de promedios entre distintos tratamientos se basa en el cociente de la variabilidad entre grupos, o tratamientos, sobre la variabilidad dentro de estos grupos. Esta forma de analizar resultados se resume en la conocida tabla de Análisis de Varianza, que todo Psicólogo ha encontrado en su carrera en algún momento.

Pero, antes de 1940, en la Psicología Americana no existía esta uniformidad en el método de análisis. Al contrario, según mencionan Gigerenzer & Murray (1987), los psicólogos utilizaban fundamentalmente datos de casos individuales que eran extensivamente descritos. El proceso de inferencia se hacía en forma fragmentada y negociando con el lector la verosimilitud de las conclusiones. Los pocos mecanismos formales para comparar promedios estaban basados en las llamadas «razones críticas», donde las diferencias entre grupos o tratamientos eran expresadas como cocientes de la diferencia entre promedios dividida por la desviación estándar de la diferencia.

Definitivamente, la introducción del Análisis de Varianza Fisheriano en Psicología representó un gran avance para la sistematización de la Psicología Experimental. Pero, ¿qué tipo de modelo de inferencia estadística fue adoptado dentro de la Psicología? y ¿por qué el Análisis de Varianza se convirtió en Psicología en *el* método por excelencia para realizar inferencias de los datos a las hipótesis?

Gigerenzer (Gigerenzer & Murray, 1970) argumenta que el paradigma de inferencia estadística que fue popularizado por los textos de Psicología, fue un

modelo híbrido que combinaba, sin poder reconciliar del todo, los enfoques de Fisher y de Neyman & Pearson. La idea Fisheriana de rechazar la hipótesis nula se mantiene como uno de los ejes centrales del proceso de inferencia, pero se admite el concepto del error de tipo II de Neyman, aunque éste no puede ser determinado dentro de una situación experimental. En los libros de texto, hallamos que la idea del error tipo II ha sido incluida en recientes ediciones como evaluación del poder estadístico de una prueba. La evaluación del poder de la prueba ayuda a determinar el tamaño de muestra necesario para realizar el experimento.

Este modelo híbrido, en el cual se puede decir que la teoría de Neyman & Pearson es el super-ego, pero que la teoría Fisheriana es el ego, no fue presentado como lo que es, es decir, un compuesto de dos modelos de inferencia estadística que están en cierta forma contrapuestos. Por el contrario los libros de textos, comenzando por el texto de Snedecor (1937), que fue el primero en difundir el Análisis de Varianza de Fisher entre los científicos sociales, presentaban este modelo híbrido como un modelo monolítico, como el método de inferencia estadístico por excelencia.

La adopción en Psicología del modelo híbrido como casi el único instrumento válido para realizar inferencias estadísticas tuvo varias razones. Danziger (1987) menciona que, la adopción de un método para evaluar diferencia de promedios entre grupos era particularmente atractiva para aquellos Psicólogos que habían derivado a un enfoque neo-Galtoniano de la investigación psicológica.

La Psicología experimental se inició con el paradigma Wundtiano, donde se recolectaba extensos datos de *sujetos individuales* y se reportaban descriptivamente las manipulaciones experimentales con miras a establecer una causalidad psicológica entre las condiciones experimentales a que *un sujeto individual* era sometido, y las respuestas que éste producía. Este programa respondía a la idea de que la causalidad psicológica tenía que ser demostrada dentro de un mismo individuo, ya que el objetivo de la Psicología era describir el fenómeno psíquico dentro de la mente de una persona. La variabilidad existente en los resultados del mismo individuo, se consideraba simplemente como un error debido a factores no controlables o aleatorios, que era eliminada a través de alguna forma de promediar los datos del mismo sujeto.

La contraparte del paradigma de Wundt, fue el paradigma Galtoniano, en el cual el interés estaba en la descripción de relaciones entre variables tal y como se daban en situaciones reales, en grupos naturalmente formados y que no eran influenciados por ningún tipo de tratamiento experimental. El énfasis estaba en describir relaciones funcionales entre variables, y no en establecer relaciones

causales entre estados mentales internos a un individuo. Metodológicamente, esta orientación se basaba en la descripción de las relaciones entre una gran cantidad de mediciones aplicadas a grandes muestras de individuos. La variabilidad entre los sujetos se consideraba no como un error, sino como aquellas *diferencias individuales* que necesitaban ser explicadas. Pronto, los Psicólogos se dieron cuenta de que, en vez de describir grupos naturales, podían construir grupos artificiales que eran factibles de someter a condiciones experimentales de laboratorio, y que a diferencia del enfoque Wundtniano, era más fácil evaluar el cambio *promedio* de los grupos que habían recibido diferentes condiciones experimentales. Hubo un cambio fundamental en esta forma de investigación: si bien la causalidad psicológica se intentaba mantener, la forma de evaluación ya no era dentro del mismo individuo, sino dentro de un *individuo promedio* definido por el proceso de promediación de los valores grupales.

Esta forma de hacer investigación, manteniendo el deseo de establecer relaciones causa-efecto, y a la vez trabajar con promedios grupales, predispuso a los psicólogos a adoptar el Análisis de Varianza de Fisher. Sin embargo, en esta adopción, el modelo híbrido fue el que llegó a adoptarse a la larga.

Gigerenzer & Murray (1987) añaden dos razones más por la cual el Análisis de Varianza de Fisher fue tan rápidamente aceptado. Primero, al convertir a la estadística inferencial como el método por excelencia para la inferencia, los psicólogos lograban una unificación de la disciplina a nivel metodológico que había sido, y es aún, muy evasiva a nivel teórico; y a la vez, esta adopción brindaba el aura de disciplina cuantitativa a la Psicología. Otra razón es que el análisis de varianza permitía a los psicólogos aplicados, especialmente los que trabajaban en Educación y Psicometría, resolver urgentes necesidades de análisis grupales.

### ***LA INFERENCIA ESTADÍSTICA COMO METÁFORA PARA LA TEORIZACIÓN PSICOLÓGICA: LA DETECCIÓN Y DISCRIMINACIÓN DE ESTÍMULOS***

Si bien la «revolución de la Inferencia» sucedió en la Psicología Americana en los cuarenta, fue con el advenimiento de la «revolución cognitiva» en los sesentas, cuando los términos estadísticos -términos como probabilidades, probabilidades *apriori*, procesos de muestreo, criterios de decisión, ratios de verosimilitud, etc.- brindaron a las teorías cognitivas un nuevo vocabulario conceptual. La esperanza era de que las herramientas estadísticas que unificaron la Psicología metodológicamente, pudiesen también unificarla teóricamente. Estas metáforas

estadísticas transformaron nuestro entendimiento de los procesos cognitivos, creando nuevas preguntas de investigación, generando nuevos datos, pero también repitiendo algunas de las confusiones en la aplicación de un modelo híbrido de inferencia estadística. La aplicación de estos conceptos estadísticos definieron la metáfora de la mente como un estadístico intuitivo. Revisaremos someramente, como la metáfora estadística cambió la forma de conceptualizar uno de los más antiguos problemas de la Psicofísica tradicional: la detección y discriminación de estímulos.

Un proceso cognitivo elemental es el detectar la presencia de un objeto, o de discriminar entre dos objetos (que puede ser conceptualizada como detectar la diferencia entre dos objetos). Por ejemplo, la detección de una señal en una pantalla de radar, discriminar si dos objetos tienen igual peso, detectar un sonido casi audible, etc. son ejemplos de un proceso de detección de estímulos. Este proceso ha sido desde hace mucho tiempo descrito en términos de la metáfora de los umbrales: Se considera que la detección de un estímulo ocurre sólo si el efecto que el estímulo tiene sobre el sistema nervioso excede un cierto valor fijo de *umbral*. La detección de la presencia de un estímulo sólo puede ocurrir si la excitación excede un *umbral absoluto*; mientras que la detección de una diferencia entre dos estímulos ocurre si la excitación producida por uno de ellos excede la producida por el otro en una cantidad mayor a un *umbral diferencial*.

Los umbrales eran medidos en unidades físicas y su relación con la intensidad del estímulo estaba expresada por la ley de Weber, la cual dice que: la «diferencia apenas notable» (*just noticeable difference*) entre estímulos, el umbral, es una fracción constante de la intensidad del estímulo. Sin embargo, en las mediciones de umbrales, los experimentadores se encontraban con una considerable variabilidad en la detección de un estímulo, aún cuando la misma intensidad física era utilizada. Esto forzó a que los umbrales fueran definidos como la diferencia en intensidad física que era detectada en un cierto, y arbitrario, porcentaje de veces.

Esta variabilidad en la detección de estímulos, sugirió la idea de que la detección de un objeto era función del nivel interno de «actividad cerebral» (Solomon, 1900). Este supuesto, que la actividad cerebral puede fluctuar, proveyó de una explicación alternativa a la teoría de los umbrales perceptuales. Fue Thurstone (1927) quien elaboró esta idea en términos de su teoría probabilística de la variabilidad discriminativa.

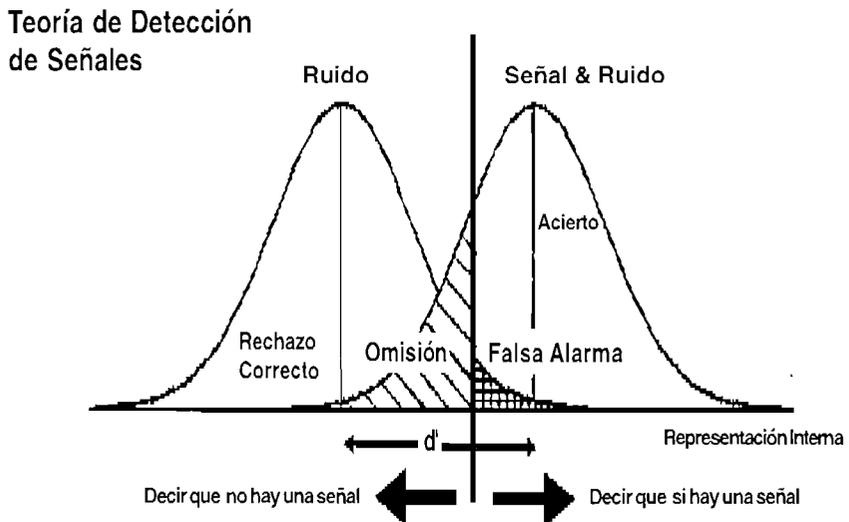
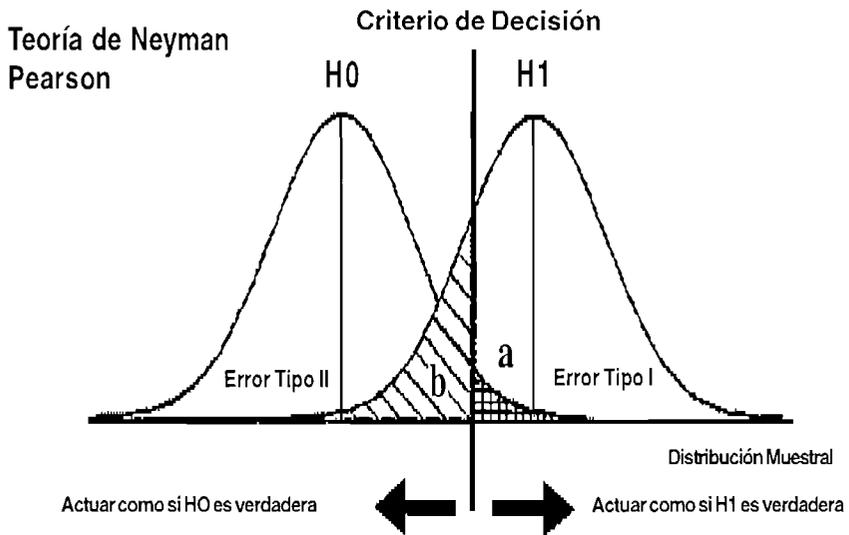
Thurstone asumió que las presentaciones repetidas del mismo estímulo físico resultaba en una distribución normal de valores subjetivos generados por

algo que él denominó un «proceso discriminativo». El promedio de esa distribución normal se define como el valor subjetivo del estímulo que uno está buscando. Cuando comparamos dos estímulos tendríamos, entonces, dos distribuciones normales sobre una misma dimensión subjetiva. El problema es como obtener los promedios de esas dos distribuciones. La solución que Thurstone propuso fue la de presentar repetidamente a un sujeto pares de estímulos y pedirle que indicara cual de los estímulos era mayor (o más pesado, más brillante, más sonoro, etc.). Es obvio, que bajo el supuesto de las dos distribuciones normales con igual varianzas, la probabilidad de decir que un estímulo es mayor que otro, depende de la diferencia entre los promedios de las dos distribuciones. Así, si ambas distribuciones tienen el mismo promedio, la probabilidad de decir que un estímulo es mayor que otro es básicamente 5. Asumiendo iguales varianzas, la probabilidad de que un estímulo A sea mayor que un estímulo B fue definida como una función de la distribución normal acumulada:  $F(a-b)$ , donde  $F$  es la función normal acumulada, y  $a$ ,  $b$ , son los promedios de las dos distribuciones normales. Esta regla se conoce como la *Ley de los Juicios Comparativos*, específicamente el Caso V de esa Ley (Thurstone, 1927).

La teoría probabilística de Thurstone tiene una estructura similar a la Teoría de Detección de Señales, que está basada en la teoría de decisión estadística de Neyman & Pearson. Pero Thurstone no le dio ninguna interpretación en términos de un proceso de estadística inferencial, o en términos del individuo como un procesador estadístico de información. Esta reformulación del proceso de detección de estímulos como un proceso de inferencia estadística se difundió en Psicología después que los métodos de estadística inferencial se popularizaron entre los investigadores en Percepción.

A mediados de los cincuenta, la teoría de prueba de hipótesis estadística de Neyman & Pearson se reformuló como una teoría de detección y discriminación de estímulos. Esta nueva reformulación se denominó la *Teoría de Detección de Señales* (Tanner & Sweets, 1954). Esta teoría se originó en el contexto de problemas de ingeniería relacionados con el desempeño de sensores ideales en tareas tales como detectar una señal en una pantalla de radar. De ahí que muchos términos, como «señal» y «ruido» tengan connotaciones ingenieriles. Tanner & Sweets importaron la analogía entre prueba de hipótesis y sensores ideales a la Psicología, modificándola en una analogía entre la prueba de hipótesis estadística y la forma en que la mente trabaja para detectar un estímulo. Explícitamente Tanner compara la tarea de un observador humano en un proceso de detección como una tarea de probar una hipótesis estadística.

La figura adjunta (Gigerenzer & Murray, 1987, pg. 43) muestra el isomorfismo entre la teoría de Neyman & Pearson y la teoría de la detección de señales.



Como mencionamos antes, en el modelo de inferencia estadística de Neyman & Pearson distinguimos dos hipótesis: la hipótesis nula  $H_0$ , y la hipótesis alterna,  $H_1$ . Asumiendo que uno u otra son verdaderas, obtenemos dos distribuciones muestrales de un estadístico muestral, por ejemplo de los promedios de la muestra. Pero, recordemos que en este enfoque, la inferencia de los datos a las hipótesis involucra *tomar una decisión*, o establecer un *criterio de decisión*, que guíe la selección de una alternativa. Este criterio de decisión es definido por el experimentador en base a un conocimiento que va más allá de la estadística. Especialmente, el experimentador podría considerar los posibles costos de los dos tipos de error, y modificar la posición del criterio de decisión para tratar de minimizar uno o el otro. Si el estadístico muestral cae a la derecha del criterio de decisión actuamos como si la hipótesis alternativa es verdadera, mientras que si el valor del estadístico muestral cae a la izquierda del criterio de decisión decimos que la hipótesis nula es verdadera.

Esta teoría de prueba de hipótesis es reinterpretada de una forma directa, y presentada como una teoría de cómo la mente detecta una señal débil sobre un trasfondo de ruido. Como Thurstone propuso, la distribución muestral bajo la hipótesis nula, se reinterpreta como la distribución de una cierta representación interna (actividad cerebral, nivel de excitación cortical, etc.), cuando sólo se presenta ruido (i.e. no señal) al individuo. De forma similar, la distribución muestral bajo la hipótesis alternativa se reinterpreta como la distribución de la representación interna cuando la señal está presente. Al igual que Thurstone, se considera que no hay necesidad de clarificar la naturaleza de la representación interna (Swets, Tanner & Birdsall, 1964), aunque se sugiere que esta es alguna medida de actividad neuronal, tal vez impulso arribando a un punto en la corteza en un determinado lapso.

Para completar la analogía, la probabilidad de un error del tipo I se reformula como la proporción de «falsas alarmas», cuando decimos que hay una señal cuando en realidad no la hay. La probabilidad de un error tipo II, se denomina la proporción de «omisiones (*miss*)», cuando decimos que no hay señal, cuando en realidad la hay. Las otras posibilidades, de realizar correctamente la decisión, se denomina «rechazo correcto» y «acierto (*hit*)». Como podemos observar, el solapamiento de las curvas depende de la distancia entre los promedios de ambas distribuciones, que se denomina *d-prima*. *d-prima* mide la fuerza de la señal: contra más grande *d-prima*, más fuerte es la señal, más separadas estarían las dos curvas, y por lo tanto más fácil de detectar la señal. A más pequeña es *d-prima*, más cercanas están las dos curvas, y por lo tanto más difícil de detectar la señal.

Pero, aún es necesario determinar como un individuo determina el criterio de decisión a seguir para enunciar si hay una señal o no. Al igual que en el caso de la teoría de Neyman & Pearson, la Teoría de la Detección de Señales, postula que el criterio de decisión está condicionado por los costos diferenciales de cometer una «omisión» o una «falsa alarma». Es más, la teoría considera que un observador es un «óptimo» tomador de decisiones si es que opta por un criterio de decisión que minimice las dos probabilidades de error, del tipo I y II (Falsas alarmas y omisiones).

En conclusión, el proceso cognitivo de detectar una señal contra un trasfondo de ruido, que antes se había conceptualizado en términos de la existencia de umbrales fijos, es ahora considerada como una decisión entre dos distribuciones normales. La mente de un individuo tomaría la decisión como si fuera un estadístico que sigue la teoría de Neyman & Pearson.

El efecto de la metáfora de la metáfora estadística fue más allá de un simple cambio en lenguaje. Generó nuevas preguntas y nuevos tipos de experimentos, nuevos tipos de datos y nuevas respuestas para viejas preguntas. Por ejemplo, la nueva analogía entre toma de decisiones estadísticas y detección de señales separó procesos sensoriales (la sensibilidad a la detección de la diferencia entre señal y ruido, que es función de la  $d$ -prima), de procesos no-sensoriales (el criterio de decisión). Si el proceso sensorial es independiente del no-sensorial, una nueva pregunta surgió, ¿cuáles son los factores que influyen la formación de un criterio de decisión? Esto llevó a su vez a nuevos tipos de experimentos, donde se incluían pagos monetarios por cada decisión correcta, y multas por cada decisión incorrecta, con la finalidad de inducir un cambio en el criterio de decisión del individuo. A su vez, un nuevo tipo de datos fueron recolectados en el pasado en experimentos de detección sólo se registraba el número de correctas y no-correctas decisiones; ahora se vio importancia de obtener información acerca de los dos tipos de errores, las omisiones y las falsas alarmas.

### ***EL LÍMITE DE LOS MÉTODOS Y LAS METÁFORAS***

Hemos visto que en Psicología los métodos estadísticos inferenciales han monopolizado el proceso de inferencia de los datos a las hipótesis. Esta institucionalización de la estadística inferencial es continuamente reforzada a través de las políticas de publicaciones de los editores de revistas especializadas, y a través del programa de entrenamiento en la mayoría de Escuelas Graduadas de Psicología en los Estados Unidos. Si bien, la adopción de la

estadística inferencial ha resultado en general beneficiosa para la Psicología, el excesivo uso de ella ha llevado también a ciertos hábitos que tienen inquietantes posibles consecuencias a nivel de la calidad de la producción científica en Psicología.

En primer lugar, el exagerado énfasis en el análisis de los datos, puede llevar a descuidar otro gran problema en la investigación Psicológica: la medición. La mayoría de las técnicas estadísticas tratan todos los datos como si estos tuvieran las propiedades típicas de los números reales. En Psicología, en cambio, es muy raro encontrar mediciones que puedan considerarse a ese nivel de isomorfismo con los números reales.

En segundo lugar, la mecanización de la inferencia, ha fomentado una enorme producción de investigación empírica que parece haber sido diseñada más con el análisis estadístico en mente, que con el problema de investigación como primera consideración. Es fácil hallar investigadores que, confundiendo lo que son hipótesis estadísticas e hipótesis teóricas, presentan sus teorías en términos de hipótesis nulas.

En tercer lugar, el hecho de que la estadística inferencial se considera en Psicología como una herramienta desprovista de contenido y que es aplicable a cualquier tipo de problema y situación, y el advenimiento de enorme y barato poder computacional, ha llevado a un extrañísimo fenómeno en muchas partes del mundo: investigadores que usan estadísticas, pero que desconocen por completo que son estas. Frente a la presión de las Escuelas de Graduados, que solicitan investigaciones empíricas, que incluyan análisis estadísticos de los datos, es enormemente común, especialmente en Educación y Psicología, hallar candidatos a PhD que contratan otros estudiantes graduados, que tienen más gusto por los números, para que les escriban el famoso «Capítulo III» (Análisis y Resultados) de sus tesis doctorales.

En cuarto lugar, la inferencia estadística tradicional se basa en comparación de promedios de grupos independientes. La relación causal entre un tratamiento experimental y una respuesta está dada a nivel de los promedios de un grupo. Este hecho fue criticado duramente por el conductismo operante (Sidman, 1960), ya que estos consideraban que la Psicología debía ser el estudio del comportamiento del individuo, no el estudio del comportamiento de un individuo *promedio*. Esta crítica es aún válida, ya que al promediar sobre sujetos, perdemos cualquier posibilidad de describir fluctuaciones dentro de los propios individuos y el efecto de los tratamientos dentro de los

individuos. El problema de extender los procedimientos tradicionales de estadística inferencial al análisis de muchas mediciones del mismo individuo a través de diferentes condiciones se centra fundamentalmente en la violación del supuesto de independencia entre observaciones. Las formas de compensar por esta violación fuerzan a asumir otras características en los datos, que son a su vez difíciles de comprobar. En conclusión, esta limitación técnica de los métodos inferenciales, también produce una limitación en la difusión de métodos alternativos de recolectar información.

Por otro lado, la metáfora estadística continua siendo una de las principales armas teóricas para modelar procesos cognitivos. Las mismas ideas del modelo de Neyman & Pearson han sido aplicadas a las áreas de la percepción (Álgebra cognitiva), memoria (búsqueda de información en memoria), evaluación de probabilidades subjetivas (Mente Bayesiana), y Teorías de utilidad subjetiva.

### *CONCLUSIÓN*

Hemos visto como una herramienta metodológica, la estadística inferencial, ha ayudado a unificar metodológicamente el tan diverso campo de la Psicología, tanto aplicada como experimental. Esta unificación, ha sido lograda a través de una presentación monolítica de una teoría híbrida de la prueba estadística de hipótesis. Pero, de simple herramienta, la estadística inferencial, especialmente la teoría de Neyman & Pearson, se convirtió en una metáfora para describir procesos cognitivos como el de la simple detección de estímulos.

El balance de la adopción de este modelo de inferencia estadística en Psicología es en general positivo. Pero, esto no implica que la particular adopción de este modo de inferencia de datos a hipótesis haya potencializado todos los avenidas de investigación de igual forma. Hoy en día, dentro de la Psicología Académica contemporánea, es difícil pensar una forma alternativa a la estadística inferencial para analizar datos experimentales. Nuevas alternativas de análisis provendrán de nuevos impactos tecnológicos, por ejemplo, del hecho de que el poder de computo ha disminuido en costo de forma asombrosa en los últimos diez años. Cualquier procedimiento alternativo de análisis de resultados -como por ejemplo las pruebas de aleatorización, (Edgington, 1987), donde se calculan todas las posibles re-agrupaciones de los datos, y se computan todos los posibles estadísticos de prueba de esas re-agrupaciones- serán basadas en el poder de computo de las computadoras personales de hoy en día.

Conceptualmente, las estadística inferencial, y en general, la teoría de probabilidades, seguirán brindando metáforas en las cuales la mente de un individuo será vista como un «estadístico intuitivo». □

## **BIBLIOGRAFÍA**

- ATTNEAVE, F.  
1959 *Applications of Information Theory to Psychology*. New York: Holt, Rinehart & Winston.
- COWLES, M.  
1989 *Statistics in Psychology: An historical perspective*. Hillsdale, NJ: Lawrence Erlbaum.
- DANZINGER, K.  
1987 Statistical method and the historical development of research practice in American psychology. En L.Kruger, G.Gigerenzer & M.S.Morgan (Eds.). *The Probabilistic Revolution: Volumen II, Ideas in the Sciences*. Cambridge, MA: The MIT Press.
- EDGINGTON, E.S.  
1987 *Randomization tests*. New York: Marcel Dekker.
- FISHER, R.A.  
1925 *Statistical Methods for Research Workers*. Edinburgh: Oliver & Boyd.
- GIGERENZER, G.  
1987 Probabilistic thinking and the fight against subjectivity. En L.Kruger, G.Gigerenzer & M.S.Morgan (Eds.). *The Probabilistic Revolution: Volumen II, Ideas in the Sciences*. Cambridge, MA: The MIT Press.
- GIGERENZER, G.; SWITJINK, Z.; PORTER, T.; DATON, L.; BEATTY, J. & KRUGER, L.  
1989 *The Empire of Chance: How probability changed science and everyday life*. Cambridge: Cambridge University Press.
- GIGERENZER, G. & MURRAY, D.J.  
1987 *Cognition as Intuitive Statistics*. Hillsdale, NJ: Lawrence Erlbaum.
- HACKING, I.  
1975 *The emergence of probability*. Cambridge: Cambridge University Press.

## LA PSICOLOGÍA Y EL ESTADÍSTICO INTUITIVO

- NEYMAN, J. & PEARSON, E.S.  
1928a On the use and interpretation of certain test criteria por purposes of  
statistical inference: Part I. *Biometrika*, 20A: 175-240.
- 1928b On the use and interpretation of certain test criteria por purposes of  
statistical inference: Part II. *Biometrika*, 20A: 263-294.
- NEWELL, A. & SIMON, H.  
1972 *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.
- SIDMAN, M.  
1960 *Tactics of Scientific Research*. New York: Basic Books.
- SNEDECOR, G.W.  
1937 *Statistical Methods*. Ames, IA: Collegiate Press.
- SOLOMONS, L.M.  
1900 A new explanation of Weber's Law. *Psychological Review*, 7:  
234-240.
- SWETS, J.A.; TANNER, W.D. & BIRDSALL, T.G.  
1964 Decision process in perception. En J.A.Swets (Ed.) *Signal detec-  
tion and recognition by human observers*. New York: Wiley.
- TANNER, W.P.Jr. & SWETS, J.A.  
1954 A decision-making theory of visual detection. *Psychological Re-  
view*, 61: 401-409.
- THURSTONE, L.L.  
1927 A law of comparative judgment. *Psychological Review*, 34: 273-286.