

## BENIAMINO Y SU MINITOUR ESTADISTICO

Oscar Bottasso  
 IDICER (UNR-CONICET)  
 Rosario, Argentina

Quienes lo conocieron en su paso por los claustros refieren que *Beniamino Cercaprofondo* era un colega de certezas llevar; y gran adherente a esa especie de proclama por la que la toma de decisiones clínicas debía sostenerse en la mejor evidencia científica.

Afecto como era a estas cuestiones parece ser que rondaba en su mente la idea de estudiar si los niveles plasmáticos de un analito del metabolismo lipídico estaban asociados con el pronóstico de la enfermedad aterosclerótica. Y como el tránsito por un terreno que le resultaba bastante escabroso, lo tenía a maltraer decidió buscar el asesoramiento de otro arribado peninsular *capomastro dei numeri*<sup>1</sup> bastante mentado por aquellos años, don *Piero Della Sottile Verità*.

Lo que sigue es un *racconto* a partir de las notas que el mismo *Beniamino* tomara durante aquella tarde, climática y mentalmente calurosa.

Con toda la cautela que impone adentrarse en un territorio donde uno transita a tuestas, el galeno expuso el problema sin rodeos. Ante lo cual, *il professore* le exigió una mayor claridad en la formulación de la idea. Según *Beniamino* existían sobradas razones fisiopatogénicas para suponer que la presencia de niveles aumentados del metabolito guardaba relación con un pronóstico más pobre de los pacientes con enfermedad cardiovascular.

-Si vamos a medir pobreza estamos complicados, señaló Don Numérico.

-Perdón Maese me refiero a un mayor riesgo de morbimortalidad a los 5 años.

Con el ceño fruncido *il Capo* dejó en claro que la investigación biomédica era una empresa destinada a obtener una estimación válida y precisa de los alcances del efecto, el cual podría ser pequeño, inexistente, o de gran porte y significativo.

-Por eso he venido, recalcó *Beniamino*.

-Pues bien, como punto de partida lo que usted sostiene debe ser confrontado con la hipótesis nula o de no asociación...que los datos de su estudio podrían rechazar y así aceptar la existencia de tal relación.

El rostro de *Beniamino* trasuntaba una suerte de desasosiego que desembocó en un pedido de auxilio, "a decir verdad vine aquí por ayuda, pero evidentemente mi piso es bastante movedizo, le agradecería que fuera más explícito"

Mandolina de por medio su ilustrísima comenzó a puntear: "Desde su introducción a principios del siglo XX, la prueba de significación de hipótesis ha causado mucho debate. Combina las pruebas de Fisher y de la dupla integrada por Neyman y Pearson. Fisher partió del establecimiento de una hipótesis nula y el valor de P, llamado el nivel de significado. O si se quiere un indicador en cuanto a la fuerza de la evidencia contra la hipótesis nula, cuanto más bajo es el valor P, más fuerte".

-Por cierto, en las investigaciones médicas que uno lee, esto de P aparece siempre.

-Efectivamente es uno de los procedimientos estadísticos más transitados en la investigación biomédica, y a veces hasta mal utilizado. Para mi gusto se le presta menos atención al enfoque de Neyman-Pearson acerca de las dos hipótesis estadísticas, nula y alternativa.

-En mi caso esta última sería la asociación del metabolito con la enfermedad.

-Así es. *Attenzione*<sup>2</sup>, que estos señores también señalaron otros aspectos muy atendibles, la probabilidad de incurrir en dos errores, uno denominado alfa o tipo I (rechazar la hipótesis nula cuando era verdad) y el otro beta o tipo II (aceptarla cuando la alternativa era cierta), a la par del tamaño de la muestra requerido para el estudio.

-Vale decir que, si se rechaza la hipótesis nula, aceptamos la alternativa, acotó *Beniamino* con un dejo de entendimiento.

<sup>1</sup> Maestro de obras de los números

<sup>2</sup> Atención

- Giusto*<sup>3</sup>, y eso se denomina inferencia estadística.
- ¿Y en base a qué se rechaza o no la hipótesis nula?
- Apelamos a una prueba estadística, la cual nos brinda valores críticos para definir regiones de rechazo y de aceptación.
- Vuelvo de desbarrancar, se lamentó *Beniamino* mientras sacudía la cabeza.
- Vediamo un po', caro dottore*<sup>4</sup>, "Un modelo estadístico es un conjunto de suposiciones, incluidas las probabilidades de las observaciones en función de la hipótesis. Cuando combinamos observaciones empíricas, quiero decir los datos del estudio, con el modelo estadístico y efectuamos inferencias, asumimos validez a lo hallado. Al elegirse valores críticos para una prueba, dejaremos de lado la hipótesis nula si obtenemos un resultado que cae en la zona de rechazo o la admitiremos si el valor se ubica en la de aceptación. De entrada, se establece la probabilidad de cometer un error de tipo I, generalmente al 5%.
- No me queda claro el tipo este.
- Tiene que ver con esto de rechazar la hipótesis nula cuando es verdadera.
- ¿Por qué 5%?
- Una probabilidad del 5% para este tipo de error significa que, a la larga, si realizáramos un número interminable de estudios idénticos al planteado con la única posibilidad de que produzcan resultados diferentes, 5 de cada 100 pruebas de significación rechazarían la hipótesis nula; si es que esta fuera cierta y los datos tuvieran validez.
- Antes de que se derritan mis sesos ¿Cómo sabe uno si ocurrió o no este error tipo I?
- Si la prueba no rechaza la hipótesis nula, podemos decir con 100% de certeza que no hemos incurrido en esta clase de error. Si contrariamente está hipótesis fuera "abatida", sólo podemos indicar cuán probable es que lo hayamos cometido.
- Si es que la hipótesis nula fuera verdadera y el estudio muy bien realizado.
- ¡Bravo Dottore!, applausi.*
- Altro che applausi, el valor de P è un male di testa*<sup>5</sup>.
- Haya paz, véalo como la probabilidad de obtener una estimación alejada de un valor especificado, o si se quiere el valor de no efecto, vinculado a la "hipótesis nula".
- El valor de P lo calcularíamos.....
- Cuando haya recolectado todos los datos, existen distintas pruebas en función del tipo de variable.
- Digamos que cuando uno lee trabajos donde se resalta la pequeñez del valor P es para cargar las tintas y dejar en claro que se rechaza la hipótesis nula por un amplio margen.
- No le quepan dudas. Pero recuerde que el nivel de significación es un mix, donde se funden o confunden error alfa y valor de P. Es más, en la literatura inglesa la abreviatura de prueba de significación de hipótesis abreviada como NHST surge de la amalgamar NHT de Neyman-Pearson, con ST de Fisher.
- Complicada la yunta esta.
- El nivel de la prueba es la probabilidad de cometer un error tipo I y lo que se hace es cotejar el nivel preespecificado de la prueba, generalmente 0.05, con el valor de P surgido de analizar los datos. Si el valor P es menor que el nivel de la prueba, se rechaza la hipótesis nula y se indica que el resultado es "estadísticamente significativo". ¡Está claro *Madonna Santa!*
- Si, pero tenían necesidad de armar esta mezcolanza.
- Le hago un poco de historia. El valor de P surgido de las ideas de Fisher constituía un índice destinado a ayudar al investigador a determinar si se modificaban los experimentos a futuro o se fortalecía la hipótesis nula. Pero esta propuesta no contó con el beneplácito de Neyman-Pearson porque en cierta medida eso implicaba una vuelta al supuesto subyacente. Fue así como plantearon las pruebas de hipótesis como una alternativa más objetiva al valor P de Fisher, al no requerir una inferencia inductiva por parte del investigador; inaceptable para ellos como

---

<sup>3</sup> Correcto

<sup>4</sup> Veamos un poco, querido doctor

<sup>5</sup> Otra que aplausos.....es un dolor de cabeza

método científico. Y mientras ambos consideraban que su formulación mejoraba las pruebas de significación, Fisher sostenía que no era aplicable a la investigación porque a menudo, durante el curso de un estudio se descubre que las suposiciones iniciales sobre la hipótesis nula son cuestionables debido a fuentes de error inesperadas. Él creía que el uso de decisiones rígidas de rechazo/aceptación basadas en modelos formulados antes de recopilar datos era incompatible con el escenario al que se enfrenta el científico.

- ¡Como médico, *bravissimo Signore*<sup>6</sup> Fisher!

- ¡Me lo sospeché desde un principio!

- ¿Y cómo terminó la historia?

-A la postre lo pragmático y tecnológico prevaleció sobre lo filosófico. Y el valor de P llegó a ser interpretado como "el error tipo I observado", particularmente porque desde lo matemático el valor P es el nivel de error alfa más pequeño en el que uno puede justificar el rechazo de la hipótesis nula.

-Pobre Fisher.

-Se cuenta que se sintió consternado al ver su "Probabilidad de significación" absorbida en la prueba de hipótesis.

-¡Cómo me gustan los entretelones!

-*Tra noi*<sup>7</sup>, hay algunas limitaciones que a menudo se soslayan.

-¡Todo oídos!

-El valor P se calcula como la probabilidad de una región crítica de posibles observaciones bajo la hipótesis nula. Pero definir lo que eso significa en la vida real no es cosa sencilla. De hecho, la vinculación entre la teoría matemática de la probabilidad con el día a día es un tema controvertido.

-Recuerdo haber escuchado hace unos cuantos años que la estadística era una mesa del mejor roble asentada sobre la superficie del mar, por esto de la probabilidad.

-Excelente síntesis.

-Y qué explicación podría darle a un novato como yo.

-La teoría sobre esta cuestión propone un espacio de "eventos" donde probabilidad viene a ser una medida, como longitud o superficie, en dicho sitio.

-Y en la bajada a tierra cómo se digiere eso.

-Algunos han sostenido que el espacio de los eventos era el conjunto de todos los resultados posibles de tal experimento. Pero como sólo obtenemos un resultado, el tema no termina de cerrar.

- *¿Cosa facciamo*<sup>8</sup>?

-En estudios epidemiológicos donde uno conforma una muestra aleatoria a partir de la población; el espacio de los eventos viene a ser, el conjunto de todas las posibles muestras aleatorias que podrían haber sido elegidas.

-Y todos conformes.

-*Abbastanza*<sup>9</sup>.

-A riesgo de resultar tedioso, tengo otra cuestión con la que uno se topa reiteradamente.

-Ni se imagina todas las que a mí me asaltan, pero dígame a cuál se refiere.

-El intervalo de confianza.

-Una buena brújula ante las tantas tormentas numéricas, diría yo.

-Sálveme del naufragio, entonces.

- ¡*Caro Signore dei dubbi*<sup>10</sup>!, el intervalo de confianza es como una oscilación a partir de una estimación puntual, hacia un rango de posibles valores en la población sobre la que queremos sacar conclusiones. El dato brinda información acerca de la magnitud del efecto y la precisión.

---

<sup>6</sup> Muy bien señor

<sup>7</sup> Entre nosotros

<sup>8</sup> Qué hacemos

<sup>9</sup> Lo suficiente

<sup>10</sup> Querido señor de las dudas

La amplitud de ese intervalo indica en qué medida la estimación puntual se ve influenciada por chance.

*-Di nuovo nel fango della confusione<sup>11</sup>*

-Tomemos dos ejemplos de riesgo relativo. En el primero de ellos el valor es de 4.8 con un intervalo que va de 1.4 a 15.0 y estadísticamente significativo ( $P < 0.025$ ). Por su parte, el resultado del segundo es 1.5, su intervalo de 0.8-2.3, y el valor de  $P = 0.15$ . A pesar de ser significativo, el primer caso da la pauta que la chance es superior a juzgar por la amplitud del intervalo y por ende menos confiable que el segundo cuya oscilación es más reducida.

- ¡Bravo Egregio Professore!, pero si no es demasiada molestia querría solicitarle una última reflexión respecto a la validez del estudio.

-Por cierto, cuanto menos válido más engañosa la prueba de hipótesis. Una investigación debe satisfacer varios requisitos y hay cuatro que siempre están en el podio. El tamaño muestral, adecuación del enfoque causal y el modelo estadístico, como así también ausencia de sesgo y factores de confusión.

*-È molto complesso per carità<sup>12</sup>.*

*-Ma no, è soltanto un lavoro che viene giocato seriamente<sup>13</sup>.* Como una suerte de estratos geológicos cada período por el que atravesó el pensamiento científico ha dejado su impronta y finalmente hemos arribado a un estatus bastante aceptable; muy asequible a esa franja de mortales entusiasmados y apasionados. Fíjese en la hoja debajo del vidrio del escritorio, es parte de un texto redactado por Battista un colega que trabaja en el piso de arriba *amico delle riflessioni in profondità<sup>14</sup>*.

- ¿De qué se trata?

-Es una síntesis acerca de las características del conocimiento científico, sin los enredos del lenguaje estadístico, por lo que no se asuste.

-No llevo conmigo los anteojos de cerca.

-Se lo leo con todo gusto: "Es crítico porque trata de distinguir certezas entre falacias a partir de una medulosa tarea racional que otorga justificación a cada tramo de conocimiento ganado. Lo cual implica adherir a métodos de investigación, ceñidos a un plan previo y posterior examen de la experiencia. Por supuesto, sin perder de vista que las técnicas de verificación evolucionan en el tiempo en busca de aprehender cada vez mejor eso que llamamos realidad. También es sistemático porque es una unidad ordenada, los nuevos conocimientos se integran al sistema, relacionándose con los que ya existían. Nada de informaciones aisladas, sino un fenomenal entramado de ideas conectadas entre sí".

-Felicite a su vecino de mi parte; y no más preguntas. Prometí pasar por el hospital para conversar sobre un paciente internado hace un par de días.

- ¿Agotado?

-Le diría que siento una especie de indigestión estadístico-matemática.

-Espero que no se sienta contrariado, o defraudado en su expectativa. Si es así lo lamento.

-Creo que son demasiadas cosas y uno queda con poca capacidad de reacción.

-Tranquílcese, le aseguro que, haciendo camino, se familiarizará rápidamente con estos nuevos protagonistas.

-Insisto en que la asimilación llevará su tiempo.

-El ejercicio mental del que hace uso para ejercer su profesión es el mismo que se requiere para investigar; de modo que no debería preocuparse demasiado. Pero independientemente de lo que usted vaya a decidir, nunca deje de sentirse aprendiz. Es la mejor actitud para seguir creciendo como persona que precede a cualquier otra cosa que encaremos en la vida.... *¡Buona fortuna Dottore!*

-Lo mismo para Ud. *Professore*, y que el próximo encuentro sea menos caluroso *in ogni modo<sup>15</sup>*.

---

<sup>11</sup> De nuevo en el barro de la confusión

<sup>12</sup> Es muy complejo por favor

<sup>13</sup> Pero no, es sólo un trabajo que se juega seriamente

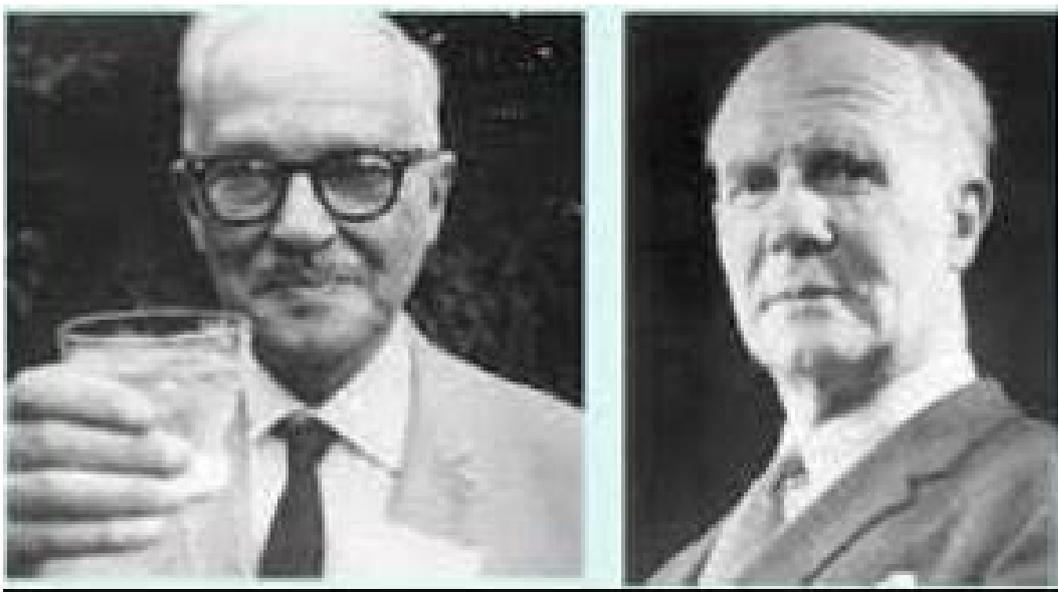
<sup>14</sup> Amigo de las reflexiones en profundidad

<sup>15</sup> En todos los modos

Las anotaciones de nuestro colega concluyen aquí y por algún avatar del destino no tan casual quedaron en uno de los cajones del escritorio que supo usar por aquellos años. Los memoriosos dicen que meses después anoticié a todos de su intención de radicarse en una mediana localidad del norte argentino donde haría las veces de médico generalista.

La afirmación de Descartes que las decisiones humanas seguían un funcionamiento similar al de un reloj no entrevió el componente emocional que se colaba en este accionar. Las adoptadas por los Médicos no son la excepción y vaya uno a saber cuántas “empatías prescriptivas” resultarían agriadas tras un análisis más detenido y meticuloso.

Nunca sabremos si algunos de estos desencantos se apoderaron de Beniamino, y si así fuera bueno sería que no les hubiera asignado gravedad. Como actividad humana la Medicina suele acarrear incertidumbres a veces desalientos y es precisamente allí cuando cobra plena vigencia aquel dicho de la Roma imperial, *Ubi dubium ibi libertas*<sup>16</sup>.



Neyman

Pearson

---

<sup>16</sup> Donde hay duda hay libertad