UNIVERSIDAD NACIONAL DE COLOMBIA

Facultad de Ciencias
Departamento de Matemáticas

Un Giro Inductivo,
a vísperas de la Perspectiva Bayesiana

28 de junio de 2022

Samuel Jacobo Garavito Segura
sgaravito@unal.edu.co

Resumen

En el siguiente artículo se hace un estudio de la metodología de Bayes desde la perspectiva de la ***Crítica Matemática*** construida a lo largo del *SFM-Critica(2022-I)*. Inicialmente, se presenta una introducción hacia la Epistemología Bayesiana, posteriormente se abarca el contexto histórico del surgimiento de la disciplina y por ultimo se presenta el análisis de factores **Críticos** propios del cimiento matemático del Bayesianismo enfocados en el cambio de modelo estadístico para el proceso inferencial.

Introducción

La necesidad de tomar decisiones ante una situación sin contemplar la suficiente información acerca de ella es uno de los grandes interrogantes de estudio que a día de hoy atañen al ser humano siendo de los mas antiguos que este ha discernido a lo largo de su desarrollo social, personal, económico, político, etc.

La predicción de un suceso ha sido un tópico sobre el cual el individuo ha posado sus ojos a lo largo del tiempo. Dominar el adelanto de un acontecimiento implicaría cierta ventaja para su beneficio frente a otros relativos a su contexto. No obstante, la mayoría de veces no se cuenta con la suficiente información para anunciar la realización de dicho suceso, en cambio, se cuenta con algunas observaciones que permiten y dan lugar a un bosquejo u aproximación del problema, pasando de considerarse un cuestionamiento predictivo a uno inferencial.

Existe una la leve diferencia que se interpone entre el concepto de *"predicción"* e *”inferencia"*. Para considerar la correcta definición del concepto de *”inferencia"*, es preciso partir desde el punto de la incertidumbre. Es válido e indispensable conocer ciertos factores influyentes y relacionados al problema de estudio en si; proponer un punto de partida permite al individuo dar un pronóstico aproximado acerca del interrogante estudiado que depende de la cantidad de información la cual este tenga a su disposición, es una relación directamente proporcional, entre mas información, mas aproximación; No obstante, al momento de contar con **TODA** la información se trasgrede el umbral de *”inferencia"* y se adentra al de *"predicción"*.

Desde, ¿Cuál será la mejor estrategia pera cazar un animal? hasta ¿Cuál será el precio del dolar el día de mañana? son interrogantes con los cuales el ser humano ha lidiado en su momento y que ha necesitado de su predicción, sin embargo, ha tenido que conformarse con sus habilidades de inferencia pues abarcar todos los factores de predicción es casi imposible debido a su carente omnisciencia.

Así nació el apotegma propio que advirtió *Thomas Bayes* cuando ideo su principio de ***”probabilidad inversa”*** en el cual estudia el problema de la determinación de la probabilidad de las causas a través de los efectos observados, algo completamente distinto al modelo probabilístico frecuentista al cual se ceñía la estadística en la época del siglo XVIII. Thomas se basa en el principio predictivo, inherente a las matemáticas, para fundar una base sólida sobre la cual soportar sus cálculos y lograr una aproximación inferencial, inherente a la estadística. Bayes planteo un nuevo modelo controversial para muchos, intrigante para otros. *Pierre-Simon Laplace, Harold Jeffreys, Richard Threlkeld Cox, Claude Shannon, etc* pertenecieron al grupo de los otros que vieron en la propuesta de Bayes un gran potencial de conocimiento, impulsándola a tal punto de que a día de hoy sea motor de un sinnúmero de aplicaciones dentro del campo de la ***Inferencia Bayesiana***, diseño de experimentos y modelación que a su vez tienen una gran repercusión sobre la toma de decisiones en casi cualquier campo del saber.

El suceso del giro Bayesiano se ha convertido en un motivo de discusión que vale la pena estudiar bajo el arquetipo de la **Critica Matemática** en aras de determinar si su filosofía epistemológica encaja correctamente en los conceptos de revolución, integración, derivación y crítica.

Contexto Histórico

Desde antes de la aparición de las matemáticas, tratar la inferencia de un suceso era un problema trabajado intuitivamente mediante un tipo de razonamiento plausible en el que se carece de la información necesaria para hacer el razonamiento formal deductivo. En el mundo real, se necesita algún tipo de extensión de la de la lógica formal.

A nivel intuitivo, el ser humano se ha vuelto bastante bueno en esta lógica extendida, y bastante sistemática. Antes de la toma de decisiones, la intuición del individuo organiza el razonamiento preliminar en las siguientes etapas.

1. Prever todas las posibilidades que pueden surgir.
2. Juzgar la probabilidad de cada una.
3. Juzgar la probabilidad de cada una de ellas, basándose en toda experiencia vivida.
4. Juzgar las consecuencias probables de las distintas acciones
5. Tomar una decisión.

Un proceso que conlleva un razonamiento crítico el cual trato a profundidad Heterodoto en el año 500 A.C. ***Heterodoto*** , historiador de la antigua Grecia, solía dar su punto de vista acerca de las decisiones políticas de los reyes persas en plena época de las llamadas *Guerras Médicas*. Afirmaba que una decisión era sabia, aunque llevara a consecuencias desastrosas, si la evidencia a mano indicaba que era la mejor; y que una decisión era tonta, aunque llevara a las consecuencias más felices posibles, si no era razonable esperar esas consecuencias.

Como ha podido notar, este tipo de razonamiento se remonta a tiempos remotos, incluso los animales ejecutan un proceso intuitivo similar que a decir verdad es el mismo desempeñado por el humano con la única diferencia tocante a su condición de ser racional. Esta pauta está tan bien organizada en nuestras mentes de forma cualitativa que parece obvio que las etapas de razonamiento anteriores pueden reproducirse en forma cuantitativa mediante un modelo matemático, de modo que la lógica ampliada sería muy útil en áreas como la ciencia, la ingeniería y la economía.

Dos siglos después, en 1713, ***James Bernoulli*** publico su nueva propuesta *"El arte de conjeturar"* donde formaliza el razonamiento lógico de la toma de decisiones. *Bernoulli* creo un modelo por medio del cual plasmar un conocimiento incompleto, el cual definió de la siguiente manera.

Definición 1.0: *Un conjunto de eventos $x_1, x_2, x_3, \dots, x_N$ con $N \in \mathbb{N}$ se define como el espacio hipotético H_0. Sea A una proposición de interés verdadera para algún subconjunto $H_0(A) \subset H_0$ y falsa para $H_0(A)^c$. Se define $|H_0(A)| = M$ la multiplicidad de A como la cantidad de casos para los cuales esta es verdadera. Por último, se define la probabilidad de A como la razón $P(A) = \dfrac{M}{N}$.*

La idea de definir este espacio hipotético radica en poner sobre la mesa aquello conocido de la manera más completa posible, es decir, recopilar toda la información a disposición. En este orden de ideas, es necesario presentar el segundo nuevo concepto.

Definición 2.0: *Dado un H_0 se define H_N su* ***espacio de hipótesis natural.***

Hacer la mención del espacio de hipótesis natural es fundamental, puesto que corresponde a las posibilidades ocurrentes en la **Natura** *(en la realidad)*, relacionadas con H_0; este conjunto de posibilidades es desconocido para el individuo, es por esto que se busca mediante la inferencia lograr algo como $H_0(A) \approx H_N$ mediante el siguiente proceso.

Inicialmente, el individuo plantea un espacio H_0 sobre el cual inmediatamente H_N está definido. Como punto de partida, el espacio H_0 y sus eventos se rechazan, son inválidos siempre que no se haya demostrado la veracidad de H_N, aunque H_N sea algo desconocido. Se supone un conjunto de proposiciones verdaderas generadas por la naturaleza $\mathscr{A} = \{A_1, A_2, \dots, A_k, \dots A_m\}$ con $m \in \mathbb{N}$ tal que A_k es la proposición más probable que fuera cierta en H_0 y se confirma por observación su veracidad, dicha sentencia se determina de la siguiente manera.

Definición 3.0: *Sea H_0 un espacio hipotético y A una proposición. Se tienen las siguientes afirmaciones:*

- *Si $|H_0(A)| \leq |H_0(A)^c|$ entonces H_0 es menos probable.*
- *Si $|H_0(A)| > |H_0(A)^c|$ entonces H_0 es más probable.*

Confirmada la observación, eso no demuestra que H_0 represente correctamente H_N, sino que no es lo suficientemente diferente de H_N para afectar la inferencia, es decir, no hay evidencia suficiente para rechazar H_0 **hasta el momento**. Si este comportamiento continúa para más proposiciones del conjunto $\mathscr{A}$, al final se obtendrá un alto nivel de confianza para el espacio hipotético H_0, de modo que no significa que sea **VERDADERO**, sino que tiene un alto valor para el proceso inferencial.

Desde otra perspectiva, dado $\mathscr{A}$ de tal forma que A_k resulta menos probable en H_0, entonces H_N es de naturaleza diferente de H_0 planteado; así, de la misma manera que fue expuesto en el ejemplo positivo, si este comportamiento continúa para más proposiciones del conjunto $\mathscr{A}$, la inferencia resultará en algo completamente erróneo. Ante un escenario así, es deber del individuo replantear H_0 basado en $\mathscr{A}$ para obtener un nuevo H_0' corregido donde no haya evidencia suficiente para su rechazo. De continuar el comportamiento negativo, se debe replantear sucesivamente $H_0^{(n)}$ veces hasta dar con un buen espacio, a fin de cuentas de eso trata el método científico **[1]**.

300 años antes, *Bernoulli* proporciono con una idea básica, sencilla e intuitiva; un gran atisbo de una de las herramientas más potentes de la estadística, las pruebas de hipótesis, que junto a ellas vienen un sinnúmero de aplicaciones más que permiten el gran avance tecnológico que a día de hoy estamos evidenciando.

Siguiendo con los acontecimientos, 50 años después de la propuesta de *Bernoulli*, es publicado el ensayo *"Essay Towards Solving a Problem in the Doctrine of Chance"*, un documento bastante peculiar encontrado por *Richard Price* entre los manuscritos del matemático y religioso ***Thomas Bayes***. En el se trataba un problema probabilístico de causalidad a través de las consecuencias **[5]**, algo de lo que no había evidencia hasta ese momento; un tipo de probabilidad inductiva a la cual se le denominó en la época como ***"probabilidad inversa"***. En ese momento *Bayes* ya había fallecido, por lo que encontrar una explicación o fundamento a dicha percepción *"Alocada"* quedo descartado inmediatamente. Desgraciadamente, el trabajo del matemático tuvo poca o ninguna influencia en el desarrollo posterior del desarrollo de la teoría de la probabilidad, los matemáticos del momento se mostraban bastante escépticos y reacios ante tal idea. Parece que esa brillante concepción casi llego a desvanecer su luz, si no fuera por la labor realizada por personajes como *Pierre-Simon Laplace* quien encontró en su trabajo aplicaciones en astronomía, geodesia, meteorología, población estadística e incluso jurisprudencia. En realidad, la mayoría del conocimiento del cual se tiene certeza hoy en día relacionado a la **Estadística Bayesiana** ha sido construido y desarrollado por diferentes figuras del campo matemático y estadístico empleando el principio de *Bayes*, debido a esto se atribuye su nombre a este amplio campo del conocimiento, pues proporciono una idea extravagante y diferente fruto de una gran teoría aun en crecimiento a día de hoy.

Con esto se concluye la contextualización para el entendimiento del estudio posterior de la teoría estadística bayesiana mediante la *Crítica Matemática*.

Estadística Bayesiana

A continuación se introducirá formalmente algunos conceptos importantes de la teoría bayesiana con el propósito de retomar el problema inferencial desde una perspectiva matemática. Primeramente, se enunciará el **Teorema de Probabilidad Total** indispensable para la mención de la **Regla de Bayes**. Así:

Teorema 1.0: *Sea $A_1, A_2, \cdots$ una partición finita o numerable del espacio muestral Ω, es decir, $A_i \cap A_j = \emptyset$ para todo $i \neq j$ y $\bigcup_{i=1}^{\infty} A_i = \Omega$, tal que $P(A_i) > 0$ para todo i. Entonces, para cualquier evento $B \in \mathscr{F}$, se satisface que:*

$$P(B) = \sum_n P(B|A_n)P(A_n)$$

Para ampliar detalles sobre conceptos y la prueba de este y próximos teoremas, puede consultar la referencia **[2]**.

En pocas palabras, lo que permite este teorema es calcular la probabilidad de un suceso a partir de probabilidades condicionadas. Gracias a este teorema se da la aparición de la regla de *Bayes* la cual enuncia lo siguiente:

Teorema 2.0: *Sea $A_1, A_2, \ldots$ una partición finita o numerable de Ω con $P(A_i) > 0$ para todo i entonces se satisface que para todo $B \in \mathscr{F}$ con $P(B) > 0$ lo siguiente:*

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_j P(B|A_j)P(A_j)} \quad \text{para todo } i$$

En aras de dar una interpretación a la regla de Bayes, supóngase que los eventos $A_1, A_2, \ldots$ son todas las posibles causas, mutuamente excluyentes, de un evento B. Bajo el supuesto que el evento B ha sido observado, la fórmula de Bayes permite conocer cuál de estas causas es la más probable de haber producido el evento B [**2**].

A partir de los dos teoremas mencionados previamente se despliega toda la teoría de la *Estadística Bayesiana.* Los modelos Bayesianos primordialmente incorporan conocimiento previo para poder estimar modelos útiles dentro de un espacio muestral y de este modo poder estimar parámetros que provengan de la experiencia o de una teoría probabilística. Esta metodología Bayesiana parte de la interpretación subjetiva (H_0) de la probabilidad y tiene como punto central el *Teorema de Bayes*, que se abordara mas adelante, en búsqueda de la actualización de la incertidumbre y toma de decisiones (H_0^n) [**3**].

Con el fin de finalmente citar nuestro Teorema, es preciso mostrar la siguiente definición.

Definición 4.0: *Sea $A_1, A_2, \ldots$ una partición finita o enumerable de Ω con $P(A_i) > 0$ para todo i. Si B es un elemento de $\mathscr{F}$ con $P(B) > 0$, entonces $(P(A_n))_n$ se llama distribución **"a-priori"**, esto es antes de que suceda B, y $(P(A_n|B))_n$ se llama distribución **"a-posteriori"**, es decir después de que sucede B* [**2**].

De aquí, es fundamental recalcar el cambio de percepción en cuanto a la teoría de la probabilidad. Previamente, se trabajaba un modelo frecuentista entendido por la frecuencia relativa de un evento esperado en el largo plazo o luego de una secuencia de ensayos. Con base en esta nueva definición, nace la consideración de un evento o del parámetro de una distribución como *Variable Aleatoria* y no como cantidad fija, una variable aleatoria que se ajustara por medio de transformaciones realizadas en el proceso de la Inferencia Bayesiana bajo el planteamiento realizado por *Bernoulli* en su momento.

Ahora si, hecha esta aclaración, es posible mencionar el **Teorema de Bayes.**

Teorema 3.0: *Sea $A_1, A_2, \cdots$ una partición finita o numerable del espacio muestral Ω, es decir, $A_i \cap A_j = \emptyset$ para todo $i \neq j$ y $\bigcup_{i=1}^{\infty} A_i = \Omega$, tal que $P(A_i) > 0$ para todo i. Entonces para cualquier $B \in \mathscr{F}$ se satisface que:*

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)}$$

donde,

- $P(A_i)$ son las probabilidades ***a-priori***.
- $P(B|A_i)$ es la probabilidad de B en la hipótesis A_i.
- $P(A_i|B)$ son las probabilidades ***a-posteriori***.

Es claro que el resultado de *la Regla de Bayes* fue posterior al mismo teorema, no obstante se hace su mención en este orden con el propósito de entender mejor la transición del modelo probabilístico.

Finalmente, establecidas todas las bases, se manifiesta el campo de la ***Inferencia Bayesiana*** donde la ejecución de proceso inferencial es literal y se da modificando ligeramente el *Teorema de Bayes* de la siguiente manera:

Teorema 4.0: *Sea H_0 un espacio hipotético para el cual sus eventos forman una partición y O una observación, entonces:*

$$P(H_0|O) = \frac{P(O|H_0)P(H_0)}{\sum_i P(O|x_i)P(x_i)}$$

donde,

- $P(H_0)$ *se llama la probabilidad* ***a-priori*** *de* H_0.
- $P(O|H_0)$ *se llama la* ***función de verosimilitud*** *de O dado* H_0*. (Probabilidad de que se cumpla la observación O dado* H_0*).*
- $P(O) = \sum_i P(O|x_i)P(x_i)$ *es la* ***probabilidad marginal*** *de O. (Probabilidad de la observación de O bajo todos los eventos del espacio hipotético).*
- $P(H_0|O)$ *se llama la probabilidad* ***a-posteriori*** *de* H_0 *dado O* **[3]**.

Cabe aclarar que actualmente al espacio hipotético H_0 se le conoce como ***Hipótesis Nula***, no obstante se utilizara esta notación en son de continuar con el hilo del escrito.

La idea de este procedimiento inferencial se soporta en obtener un conjunto $\mathscr{O} = \{O_1, O_2, O_3, \cdots\}$ finito o enumerable de observaciones mediante la experimentación tal que $P(H_0|\mathscr{O}) \approx P(H_N)$. La expresión anterior solo es una representación de la idea principal, pues $P(H_N)$ no esta definido por la misma naturaleza incierta del espacio hipotético natural. En resumen, a medida que se obtenga un mayor número de observaciones, $H_0(\mathscr{O})$ sera mas o menos probable, motivo por la cual se rechazara el espacio hipotético y se construirá uno nuevo realizando un ajuste, o se mantendrá para continuar con su estudio, esto de manera simultánea al proceso de experimentación **[4]**.

Presentada la exploración formal de la metodología Bayesiana, se procederá a realizar su estudio desde la ***Crítica Matemática.***

Crítica Matemática

Como ha podido observar, el escrito posee especial enfoque en la transición del enfoque probabilístico. Un paso frecuentista a un escalón probabilístico, un paso deductivo a un escalón inductivo. El modelo presentado por *Bayes* es una formalización totalmente natural del concepto intuitivo de la toma de decisiones. Uno de sus problemas radica en la metodología de la propuesta que se soporta sobre un razonamiento inductivo. La matemática en su base lógica parte de un proceso deductivo para llegar a una conclusión especifica y es precisamente esa una de las grandes controversias que hubo su momento para aceptar esta nueva concepción. Sin embargo, cada uno de las pruebas presentadas no viola ninguna de las reglas de inferencia lógicas, por ende a final de cuentas no significaba un dilema de gran envergadura.

Por otra parte, de las controversias más sustanciales acerca del Bayesianismo yace concebir el parámetro de una distribución como una variable aleatoria donde la respuesta a esta réplica es tan simple como un error semántico cometido por los Bayesianos en sus afirmaciones. Resulta que al momento de referirse a la *"distribución del parámetro"*, se refieren a la distribución de la probabilidad de conocimiento del parámetro. Tanto distribución ***a-priori*** como ***a-posteriori*** no representan ninguna propiedad medible del parámetro, este sigue siendo desconocido; la labor de la *"distribucion"* es determinar su estado, en pocas palabras no es una medida de cálculo, sino una medida de decisión, como se ha venido resaltando [**1**].

De esta manera, el Bayesianismo resulta un modelo de percepción que aumenta la diversidad ideológica de las matemáticas, pues provee un claro umbral **trans-ticial** entre el modelo antiguo y el nuevo. Da lugar a una nueva gamma de conocimientos que enriquece tanto a nivel teórico como a nivel experimental el ámbito matemático. Asimismo, proyecta su versatilidad sobre cualquier campo del saber, que a su vez permite la expansión y desarrollo del conocimiento sobre otros también.

La metodología de Bayes no solamente es atribuida al mismo *Thomas*, la solidez y cimentación de la misma se debe a estudios previos como el de *Bernoulli* y posteriores como el de *Laplace*. Bayes en su obra brindo su grano de arena reflejando desde una perspectiva formal el camino decisivo ante un suceso, lo demás contribución viene dada por las figuras que supieron interpretarla y darle su utilidad a tal punto de considerarla una de las grandes ramas de la estadística con nombre propio: Estadística Bayesiana, muestra de un gran arquetipo de estratificación:

1. **Cimentación:** Idea intuitiva a formalizar.
2. **Construcción:** Formalización de la idea.
3. **Postulación:** Presentación de la idea a público.
4. **Interpretación:** Estudio de la idea.
5. **Contracción:** Cuestionamiento de la idea.
6. **Experimentación:** Práctica de la idea.

Es satisfactorio ver lo bien que se acomoda la **metodología Bayesiana** a la **Crítica Matemática**, pues además de demostrar un comportamiento tran-ticial y estratificado, también prueba uno residual. Tanto Bayes como aquellos quienes seguirían su propuesta vieron en la propuesta de **Heterodoto**, **Bernoulli** y seguramente de muchas figuras más un motivo por el cual plasmar formalmente la toma de decisiones. Toda matemática surge bajo una causa justa que impulsa su desarrollo, pues sujeta a un lenguaje lógico, su único fin radica en alcanzar conclusiones de manera clara y concisa.

Para terminar este análisis, es preciso mencionar como funciona el proceso de ***Integración y Derivación*** del Bayesianismo. Como ya se ha mencionado, esta metodología lograr integrar muy bien teoría y experiencia, pues entre más observaciones empíricas se consideren, más valor y fuerza cobrara el fundamento teórico, lo cual permitirá derivar en conclusiones de mayor provecho. A los ojos de la **Matemática**, el Bayesianismo parte de una idea elemental tan potente que permite trabajar tan cerca de la lógica para aparentemente encontrar un poco de certeza acerca de lo incierto.

Conclusión

En total se ha observado que la ***Estadística Bayesiana*** cumple con los estándares bajo la revisión de la ***Critica Matemática***, además de muchos otros criterios. De hecho, más allá de un teorema o una fórmula, se ha demostrado que la base de la misma se encuentra sobre la ambición del ser humano, su ambición frente a la epistemología. Se ha demostrado que la mayoría de veces solo hace falta una poco de tiempo, creatividad, astucia, intuición, intención, lápiz y un papel para proponer una idea revolucionaria. Así hemos visto que de una idea que a poco de permanecer sepultada junto a su redactor, ahora nos valemos toda la población para hacer de nuestra vida una más amena, provechosa e intuitiva.

Por otra parte, también es necesario de contar con un legado, un legado bien instruido en el área de las matemáticas, ya que aunque tal vez un individuo sea capaz de dar a luz tal idea, puede que este no de con su utilidad ni potencial, no de cuenta de ello. Motivo por el cual futuras generaciones, con un mismo propósito o incitados por el interés que genera tal aseveración, continúen con su investigación y logren obtener avances o ¿por qué no? dar con la solución. Es la promoción del conocimiento la que hace de él la divisa más valorizada.

Bibliografía

[1] Jaynes, E. T. (1996). BAYESIAN METHODS: GENERAL BACKGROUND. Cambridge CB2 1TP.

[2] Blanco, L. (2004). Probabilidad (1.a ed.). Universidad Nacional de Colombia.

[3] Mesa., L., Rivera, M., & Romero, J. (2011). Descripción general de la Inferencia Bayesiana y sus aplicaciones en los procesos de gestión. La Simulacion al Servicio de la Academia, 2.

[4] Gelman, A. (2022). The Development of Bayesian Statistics.

[5] Bayes Thomas (1763). LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. SPhil. Trans. R. Soc.53370–418