

Autor y Mini-Curriculum:

Mariano Acciardi
Licenciado en Psicología - UBA
MN 23071 MP 20165
Mail: mariano[at]acciardi.com.ar
Sitio Web: <http://www.marianoacciardi.com.ar>
Especialista en Clínica de las Psicosis
Doctorando Facultad de Psicología UBA
Docente Facultad de Psicología UBA
Investigador UBACyT

Resumen en Español:

Se enuncian los conceptos fundamentales para la aplicación de las redes bayesianas (BN) al diagnóstico psicopatológico. De los modelos-base para sistemas expertos las BN es uno de los que más se adapta aun terreno plagado de incertezas como lo es el diagnóstico. Se sientan las bases para la construcción de este tipo de red en el dominio de la psicopatología. Las BN parten de una fórmula muy simple para el cálculo de probabilidades condicionales, y generalizan esta fórmula para el trabajo en el seno de un grafo acíclico orientado. Este grafo da cuenta de “correlaciones estadísticas” entre enfermedades por un lado; síntomas y estudios por el otro. En una red orientada al diagnóstico deben operacionalizarse en primer lugar los “Nodos Hipótesis” que constituyen las patologías a diagnosticar y a luego los “Nodos Información” o evidencia sobre las cuales se confeccionan tablas probabilísticas que dan cuenta de la dependencia condicional representada por los lazos de la red. Los lazos entre los nodos pueden confeccionarse por “Prior Knowledge” o mediante al análisis directo de los datos Una vez armada la red ésta puede utilizarse como herramienta auxiliar para el diagnóstico.

Metodología: Investigación bibliográfica y ensayo con prototipo.

Abstract in English:

We enunciate the basic concepts to consider the application of Bayesian networks (BN) in the psychopathological diagnosis. BN are one of the most used model to based expert system , which will adapt even more in fields with uncertainties such as the diagnosis. In work we give

the essential foundations for building such a network in the domain of psychopathology. The BN start from a simple formula for calculating conditional probabilities, and generalize this formula to work within a directed acyclic graph. This graph gives an account of "statistical correlation" between disease on the one hand, symptoms and studies on the other. In a diagnosis-oriented network first should be operationalized the "Nodes Hypothesis" which are the pathologies to diagnose and then, in second place, "Nodes Information" or evidence on which to arm themselves probability tables for conditional dependence represented by bonds connecting the network nodes. The links between nodes may be made by "Prior Knowledge" or by direct analysis of the data network. When the network is recompiled with probabilities tables, it can be used as an auxiliary tool for diagnosis.

Methods: Bibliographic research and testing prototype.

Construcción de una Red Bayesiana para el Diagnóstico Psicopatológico

Como producto del entrecruzamiento de la práctica clínica de la psiquiatría y la psicología de principio del siglo pasado surge la psicopatología como disciplina encargada de organizar el campo de las llamadas patologías psíquicas. A pesar de las disidencias, hay cierto acuerdo en que la Psicopatología sería el intento de construir una psicología científica capaz de estudiar el campo de las perturbaciones de las funciones psíquicas. El rigor lógico de la teoría que brinda el estudio de la lógica formal y las estadísticas han sido siempre las principales herramientas auxiliares de la psicología para constituirse como científica.

En el presente trabajo se propone una utilización particular de ciertos instrumentos estadísticos ya fuertemente utilizados en otros campos como herramientas auxiliares del diagnóstico, pero específicamente el campo de la Psicopatología: Las redes Bayesianas.

Originariamente la conexión entre estos campos de aplicación y las redes mencionadas ha sido establecida por investigadores como David Heckerman (1995) y Eric Horvitz (1993). Cabe agregar que estos investigadores son dos de los principales impulsores del desarrollo en investigación que han tenido dichas redes en los últimos años. Asimismo Pearl (1986) acuña el concepto de “Bayesian Belief Network” abriendo las puertas a su aplicación en los más diversos campos en donde la incerteza sea moneda corriente, debiendo lidiar con ella para toda resolución de problemas ulterior. Específicamente en diagnóstico médico las investigaciones de Beinlich y otros (1989) son un aporte imprescindible, así como también los avances que los algoritmos de aprendizaje y refinamiento de los datos de la red que han propuesto Spiegelhalter & Lauritzen (1988) para el procesamiento paralelo de una red demasiado grande como para ser procesada secuencialmente, lo que suele ser el caso del diagnóstico de ciertas enfermedades que pueden planteadas en términos de un “Sistema Complejo” (García, 2006) en razón de la complejidad de las determinaciones que subyacen a la presencia de una u otra afección. Estos son solo algunos de los investigadores que en este momento se encuentran estudiando los diversos modos de utilización de estas redes para el diagnóstico médico-psiquiátrico.

Como antecedentes importantes a mencionar (aunque no los únicos) que en este campo se encuentran funcionando de manera exitosa tenemos: el sistema PATHFINDER, disponible hoy comercialmente como sistema experto de decisión para patólogos quirúrgicos; el sistema MUNIN, constituido por una red múltiplemente conectada par el diagnóstico de desórdenes

neuromusculares; o el sistema PROMEDAS ampliamente utilizado para diagnóstico en endocrinología.

Existe también el corto y preciso trabajo de Wang X., Qu H., Liu P y Cheng Y (2004), financiado por la National Natural Science Foundation of China, como un excelente ejemplo de diagnóstico médico y muy esclarecedor respecto de la metodología a seguir para el armado de estos modelos.

Una Red Bayesiana (BN) es un modelo gráfico que permite representar el modo en que un humano razona. El hecho del que el ser humano sea capaz de razonar e inclusive hacerlo correctamente aún sin disponer de todos los elementos que serían necesarios para realizar una inferencia o tomar una decisión correcta, es lo que motiva las investigaciones de la inteligencia artificial de los últimos años, en el diseño y desarrollo de modelos expertos basados en este tipo de tecnologías.

Una red Bayesiana se compone básicamente de “Nodos”, que representan variables generalmente discretas, cuya fuerza de relación entre ellas se cuantifica mediante una distribución de probabilidades condicionales que determinan el valor final de aquellos nodos que no se han cargado como evidencia.

La relación entre dichas variables se establece mediante “arcos” que representan gráficamente una determinación entre nodos.

Desde 1919, en estadística los autores se han basado en la teoría de la aproximación probabilística relativa creada y formalizada por Richard von Mises. La teoría de Mises, dice que si se realiza en cualquier campo la misma experiencia muchas veces, manteniendo invariantes las condiciones del experimento, la frecuencia relativa de aparición de un determinado hecho se aproxima siempre al supuesto teórico que se daría en una secuencia uniforme de los mismos hechos. Todas las técnicas de muestreo existentes se basan en tal suposición para poder realizar el salto de determinar a partir de un número finito de individuos propiedades que teóricamente deberían darse para el caso en que pudiera captarse el colectivo completo relativo a tales hechos. Sin ello la estadística sería incapaz de dar cuenta de algo que ocurra en el mundo.

Hay fragmentos de la realidad que no pueden fácilmente subsumirse bajo la aproximación de las frecuencias relativas, muchas veces por la imposibilidad de reproducir la experiencia un número n de veces manteniendo las condiciones invariantes. Por ejemplo, para un partido de

fútbol. Es imposible mantener invariantes las múltiples condiciones que pueden determinar el resultado del mismo. Qué históricamente hasta el momento un x equipo haya ganado un 30% de las veces, no nos dice casi nada respecto del resultado que obtendrá en el próximo partido. Ahora bien, dicho porcentaje puede determinar un cierto grado de creencia subjetiva respecto del próximo resultado. Las aproximaciones a las probabilidades bayesianas intentan de alguna manera determinar el grado de creencia que un determinado sujeto puede tener sobre un evento, lo que dista mucho de lo que sería la posibilidad teórica de que una moneda salga cara o seca en un número determinado de tiradas. No se trata aquí de razones matemáticas, sino de otro tipo de "ratio". No es posible pensar en lo que a probabilidades bayesianas corresponde, datos objetivos acerca de cuál es el colectivo supuesto del evento, se trata de una opinión o creencia que es siempre discutible. En este punto poco tienen por hacer las frecuencias relativas de las estadísticas tradicionales.

Es por eso que recientemente ha cobrado cada vez mayor importancia la aproximación y el uso distinto de la estadística basada en la llamada "Aproximación Bayesiana", lo que se acercaría al modo de razonar de un individuo humano.

Respecto de esto último, con Cox (1961), podemos afirmar que un gran número de inferencias subjetivas [si no todas] pueden ser formalizadas mediante un álgebra Booleana adecuada. De acuerdo a este autor, en el álgebra Booleana el significado de la proposición ocupa la función reservada a las cantidades en el álgebra tradicional.

Las reglas que gobiernan la inferencia probabilística han sido enunciadas por este autor basándose en dos principios básicos:

1-La probabilidad de inferencia dada una evidencia determina la probabilidad de su contradictoria.

2-La probabilidad dada una evidencia de que dos inferencias sean ambas verdaderas está determinada por cada una de sus probabilidades.

De esta manera, es posible utilizar la inferencia probabilística como guía y apoyo a decisiones razonables. Cox (1961) se encarga de formalizar en álgebra booleana, las reglas por las que cotidianamente juzgamos una alternativa como más probable que la otra, o que tal inferencia es tan certera que puede ser tomada por dada, o una contingencia posible, tan cercana a la imposibilidad que se puede dejar fuera de consideración.

La aproximación bayesiana a la inferencia probabilística es un modo de poner en juego esta estructura subyacente al razonamiento humano, formalizando la manera en que es posible

inferir probabilidades desconocidas a partir de un conjunto de probabilidades conocidas. El teorema de Bayes aplicado a la inferencia probabilística es aquello que puede denominarse “Estadística Bayesiana”. Esta rama de la estadística propone la utilización de dicho teorema para inferir probabilidades desconocidas partiendo de probabilidades conocidas.

El teorema de Bayes en términos generales establece que (Neapolitan, 2004):

Dados dos eventos E y F, de tal manera que la $P(E) > 0$ y la $P(F) > 0$ tenemos:

$$P(E|F) = \frac{P(F|E) P(E)}{P(F)}$$

El teorema que refleja esta fórmula fue creado en la edad media por Thomas Bayes (esta y su fórmula generalizada fueron publicadas por primera vez en 1763), sin embargo, su aplicación reciente a una red acíclica entre nodos continuos potencia enormemente su aplicabilidad y capacidades.

Se enunciarán brevemente algunos de los conceptos básicos de estadística a fin de comprender el sesgo de la aproximación bayesiana.

- a) “Espacio Probabilístico” se denomina a la conjunción de un conjunto de todos los resultados posibles de un experimento aleatorio cualquiera (Ω), una función de probabilidad que implica un grado de incertidumbre de ocurrencia entre 0 y 1 para cada uno de los subconjuntos de Ω .
- b) Una “Función Probabilística” es aquella que asignará a cada suceso que constituye la base del experimento aleatorio, un valor entre 0 y 1 que da cuenta de la posibilidad de ocurrencia de tal suceso.
- c) La “Probabilidad Condicional” es la probabilidad de que ocurra un evento dado, teniendo por conocido otro evento de esta cadena.

La estadística bayesiana se basa en este concepto, enriquecido con la teoría de grafos.

Bayes simplemente enuncia el algoritmo para su cálculo.

Desde el punto de vista matemático una red bayesiana es una representación compactada de la tabla de cálculo de probabilidad conjunta de un espacio probabilístico que compete a un sector dado de la realidad.

Una BN (Bayesian Network) no implica solo probabilidad, sino que además presenta un modelo gráfico. Intuitivamente este modelo gráfico, en lo que sería la comunicación interpersonal humana, da cuenta de una semántica muy simple denominada causalidad, semántica habitual en el hombre, mas allá de las disquisiciones filosóficas que este concepto pueda suscitar. Un segundo uso de un modelo gráfico puede ser el de la comunicación con

una computadora. En este caso, deben poder especificarse reglas muy estrictas de construcción y un lenguaje que pueda ser procesado de acuerdo a reglas bien determinadas y explícitas que permitan comunicar la semántica del gráfico de un modo traducible en operaciones de computadora.

En una BN se distinguen aspectos cualitativos y aspectos cuantitativos. Ambos deben estar bien definidos para que el modelo pueda ser objeto de un procesamiento informático.

Formalmente, el aspecto cualitativo de una BN se representa de acuerdo a la teoría de grafos, pero esta representación gráfica en una BN permite además definir las reglas cuantitativas de cálculo, que permiten construir esta “representación compacta”

El aspecto cuantitativo de una BN se representa mediante tablas de probabilidades asociadas a cada nodo. De acuerdo al tipo de nodo de que se trate, el tipo de probabilidades que serán expresadas en esta tabla son:

1) Para “Variables Hipótesis” se expresan inicialmente “Probabilidades previas”, es decir previas a toda evidencia específica: Por Ej, Existe un 2% de probabilidades, independientemente que ya lo haya entrevistado o no, que un sujeto sea un paranoico. $P(H)$

2) Para “Variables de Información” se calcula la probabilidad de ocurrencia del evento I dada la hipótesis H: $P(I|H)$. Es decir, la probabilidad de ocurrencia de que un sujeto tenga alucinaciones siendo un paranoico.

3) Para un set de observaciones f_1, f_2, \dots, f_n ... dado para los nodos de Información I_1, I_n .. (Es decir diferentes combinaciones de variables de información para un caso dado) se calcula el producto de todas las probabilidades para cada uno de esos casos: $P(f_1, \dots, f_n|H) = P(f_1|H) \times P(f_2|H) \times \dots \times P(f_n|H)$. Esto se denomina la “Probabilidad Posterior” luego que a la red se la alimenta de suficiente información como para calcularla.

Estos tres items resumen el modelo más usualmente utilizado como herramienta auxiliar para el diagnóstico médico: Naive Bayes Models.

Una vez operacionalizadas las variables, se deben establecer los arcos para luego poder definir los parámetros asociados a cada nodo/variable. Estos parámetros constituyen las tablas de probabilidad condicional de acuerdo a los lazos establecidos con sus padres.

Hay básicamente dos modos de definir los “arcos” o relaciones “causales” entre los nodos. El primero de ellos corresponde a la concreción del conocimiento disciplinario en donde un “experto” realiza la vinculación según su mejor conocimiento y parecer. Este modo de armado tiene como deficiencia el hecho de que la determinación subjetiva de relaciones de probabilidad condicional no es en todos los casos o en todos los “expertos” la misma. Por

ese motivo suele requerir muchísimos ajustes hasta tanto se torne suficientemente óptimo. A esta forma de armado del modelo se la conoce como “Prior knowledge”

El segundo corresponde a un “aprendizaje” de los datos. Para este fin se han creado diversos algoritmos que permiten, a partir del análisis de los datos y con la grilla de una operacionalización de indicadores previa, derivar la topología de la red. Este modo de inferencia tiene la ventaja de que si existe una cantidad suficiente de datos pueden inferirse uno o varios modelos que posiblemente se encuentren más próximos a una realidad concreta que si el modelo de la la topología ha sido realizado puramente a partir del “Prior Knowledge” de algunos expertos.

En síntesis, para el armado de una BN aplicada al diagnóstico patológico se deberían seguir los siguientes pasos (IMPORTANTE: Se presentarán aquí porcentajes solo a título ilustrativo del armado de la red, no necesariamente corresponden a la realidad):

I) Definir las “Variables Hipótesis” con sus respectivos porcentajes de prevalencia epidemiológica como ser: A) Paranoia: 0.30 %; B) Esquizofrenia: 2.00 % y C) Delirium Tremens 0.10 %.

II) A continuación se establece una relación de cada una de las patologías con las variables de información relevadas que constituyen los síntomas u otros efectos de las patologías definidas en (I) -resultados de análisis , etc.- : a) Delirio; b) Afección de la voluntad; d) Temperatura mayor a 38°; Etc.

III) Luego de lo cual se establecen las probabilidades condicionales respecto de cada uno de sus padres, tanto por la negativa como por la positiva, a saber:

A) La Probabilidad que haya “Delirio” cuando un sujeto se encuentra aquejado de “Paranoia” es del 80 %, que no lo haya 20%; su inversa: La Probabilidad que un sujeto que no es Paranoico presente esta sintomatología es del 2%, y que no la presente es del 98%.

B) La Probabilidad que un sujeto presente “Afección de la Voluntad” cuando se encuentra aquejado de “Esquizofrenia” es del 90%, que no la presente es del 10%. La Probabilidad que presente dicha sintomatología un sujeto que no es “Esquizofrénico” es del 30%. La Probabilidad de que no la presente es del 70%.

Deben construirse las tablas de probabilidad condicional respecto de todos los “Síntomas” o “Efectos” de acuerdo a los arcos de la red.

IV) Como cuarto paso se compila la red y la misma ya se encuentra lista para recibir las evidencias, esto es modificar los porcentajes de los nodos de las “Variables de Información” de acuerdo a cada caso específico. Automáticamente esto actualiza los porcentajes de las

“Variables Hipótesis”, las que deberían aproximarse progresivamente al resultado diagnóstico a medida que se agrega evidencia a la Red.

Referencias:

- Beinlich I. A., Suermondt, HJ, Chavez, R. M, Cooper, G. F. (1989). The ALARM monitoring system: A case study with two probabilistic inference techniques for belief networks. Proceedings of the Second European Conference on AI in Medicine Vol. 38, (pp. 247-256). London, United Kingdom: Springer Verlag.
- Bercherie, P. (1986). Fundamentos de la Clínica; Buenos Aires. Argentina: Manantial
- Cox, R. T. (1961). Probability. En *The Algebra of Probable Inference*. London, United Kingdom: Johns Hopkins University Press (USA Edition).(pp 1-29).
- García, R. (2006). Conceptos, métodos y fundamentación epistemológica de la investigación interdisciplinaria. En *Sistemas complejos*. Barcelona, España: Gedisa.
- Heckerman D. (1995). A Tutorial on Learning With Bayesian Networks. Recuperado el 12/06/2007 de <http://research.microsoft.com/pubs/69588/tr-95-06.pdf>
- Horvitz E. J. (1993). Automated Reasoning for Biology and Medicine. Recuperado el 30/05/2008 de <ftp://ftp.research.microsoft.com/pub/ejh/aibio.pdf>
- Jensen F.V. y Nielsen T.D. (2007). Bayesian Networks and Decision Graphs. New York, USA: Springer Science + Business Media. p.43.
- Lauritzen S. L. , Spiegelhalter D. J.(1988). "Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems". *Journal of the Royal Statistical Society, Series B, Vol. 50* (pp. 157-224).
- Naïm P. et autres (2007). Réseaux Bayésiens. Paris, France: Groupe Eyrolles.
- Neapolitan R.E. (2004). Learning Bayesian Networks. New Jersey, EEUU: Pearson Prentice Hall.(pp. 23-27)
- Pearl, J (1986). "Fusion, propagation, and structuring in belief networks". *Artificial Intelligence* (Elsevier) 29 (3): 241–288. doi:10.1016/0004-3702(86)90072-X. ISSN 0004-3702
- Pearl J. (1999): On Causality. Recuperado el 10/10/2009 de <http://bayes.cs.ucla.edu/BOOK-99/book-toc.html>
- Wang X., Qu H., Liu P, Cheng Y. (2004). A self-learning expert system for diagnosis in traditional Chinese medicine. *Expert Systems with Applications, Volume 26, Issue 4, May 2004, (pp 557-566), ISSN 0957-4174, doi: 10.1016/j.eswa.2003.10.004*

Notas:

i [Probabilidad de existencia del hecho F sea distinta de 0]

ii [Probabilidad de existencia del hecho F sea distinta de 0]

- i [Probabilidad de existencia del hecho E sea distinta de 0]
- ii [Probabilidad de existencia del hecho F sea distinta de 0]