

EL PENSAMIENTO PROBABILÍSTICO EN ESTUDIANTES DE INGENIERÍA EN SISTEMAS DE INFORMACIÓN

Probabilistic thinking engineering students in information systems

AMABLE MORENO*

Recibido: 3 de Noviembre de 2015. Aceptado: 19 de Noviembre de 2015

DOI: <http://dx.doi.org/10.21017/rimci.2016.v3.n5.a3>

RESUMEN

En este trabajo presentamos los resultados de un estudio realizado con ochenta y un estudiantes de la carrera Ingeniería en Sistemas de Información. El objetivo es determinar las tendencias de pensamiento probabilístico de los estudiantes, a partir de las respuestas que dieron a un cuestionario. El cuestionario está conformado por dos dimensiones; una relativa al reconocimiento de la aleatoriedad y la otra a la estimación de la probabilidad. Las respuestas de los estudiantes fueron analizadas mediante técnicas estadísticas multivariantes: análisis de clusters y análisis discriminante. Los resultados muestran la presencia de tres tendencias de pensamiento probabilístico: Tendencia al Determinismo, Tendencia a la Causalidad y Tendencia a la Incertidumbre.

Palabras clave: clusters, discriminante, estudiantes, ingeniería, pensamiento, probabilístico.

ABSTRACT

This paper presents the results of a study of eighty one students of engineering career in information systems. The objective is to determine the trends of thought students, from the answers given to a questionnaire. The questionnaire consists of two dimensions; one on the recognition of randomness and the other the estimation of the probability. The student responses were analyzed by multivariate statistical techniques: cluster analysis and discriminant analysis. The results show the presence of three trends probabilistic thinking: Determinism, Causality and Uncertainty.

Keywords: clusters, discriminating, students, engineering, thinking, probabilistic.

I. INTRODUCCIÓN

EL PRESENTE trabajo consiste en un estudio sobre las tendencias de pensamiento probabilístico que tiene un grupo de estudiantes de la carrera de Ingeniería en Sistemas de Información. Este trabajo es la continuación de un estudio más amplio realizado con los estudiantes para profesor de Biología ([1], [2]) y para profesor de Matemática ([3], [2]), de la provincia de Mendoza, Argentina; en el que

detectamos la existencia de cuatro tendencias de pensamiento probabilístico: tendencia al Determinismo, tendencia al Personalismo, tendencia a la Incertidumbre y tendencia a la Contingencia.

Entendemos que los estudiantes de tercer año de la carrera de Ingeniería en Sistemas de Información han construido esquemas cognitivos más elaborados que los estudiantes de cursos inferiores. En nuestro caso, nos referimos al esquema

* Doctora en Enseñanza de las Ciencias y la Tecnología. Licenciada en Matemáticas. Licenciada en Estadística, Universidad Tecnológica Nacional, Facultad Regional Mendoza, Departamento de Matemática. Profesora de Probabilidad y Estadística en las carreras de Ingeniería. Correo electrónico: amable.moreno@frm.utn.edu.ar

cognitivo que organiza las características asociadas al concepto en un todo integrado en la memoria, y es usado en un procesamiento cognitivo tal como recordar, reconocer, razonar y tomar decisiones. Es decir, el esquema que organiza características del fenómeno aleatorio en categorías; el mecanismo físico de la producción de resultados, las secuencias de resultados, la probabilidad, la frecuencia esperada, la variabilidad, la enumeración de los elementos del espacio muestral, las clases particulares de fenómenos aleatorios en distintos contextos.

II. DISEÑO METODOLÓGICO

En el estudio participaron 81 estudiantes de tercer año de la carrera Ingeniería en Sistemas de Información, correspondientes a dos comisiones; de la Facultad Regional Mendoza de la Universidad Tecnológica Nacional; de la provincia de Mendoza, Argentina.

Las técnicas de recogida y análisis de datos que hemos utilizado se enmarcan en el enfoque cuantitativo. El cuestionario utilizado fue tomado de Cardeñoso [4] sobre el cual se realizaron algunas modificaciones para adaptarlo a nuestro contexto físico y cultural. Las categorías usadas por el autor mencionado para el diseño del cuestionario fueron las siguientes:

Causalidad: Argumentaciones que tienen como criterio de reconocimiento de la aleatoriedad explicaciones en función de los diversos factores causales o en la ausencia de posibilidad de su control.

Multiplicidad: Argumentaciones que tienen como criterio de reconocimiento de la aleatoriedad la existencia de múltiples posibilidades en el desarrollo del fenómeno.

Incertidumbre: Argumentaciones en las que se utiliza como criterio de reconocimiento de la aleatoriedad la propia imprevisibilidad del suceso, sin profundizar en su explicación o análisis.

Subjetiva: Argumentaciones en las que utiliza como criterio de reconocimiento de la aleatoriedad consideraciones referidas a la propia vivencia o creencia subjetiva.

Contingencia: Argumentaciones estimativas de cuantificación de la probabilidad basadas en la comparación entre los casos favorables y desfavorables de un suceso.

Laplaciana: Argumentaciones estimativas de cuantificación de la probabilidad basadas en la proporción entre los casos favorables y casos posibles del fenómeno.

Frecuencial: Argumentaciones estimativas de cuantificación de la probabilidad basadas en la lectura frecuencial del fenómeno o de la información aportada.

Equiprobabilidad: Argumentaciones estimativas de cuantificación de la probabilidad basadas en justificaciones desde la equiprobabilidad entre los resultados del fenómeno.

Experiencial: Argumentaciones estimativas de cuantificación de la probabilidad basadas en criterios fruto de la experiencia personal.

Las respuestas de los estudiantes al cuestionario se transformaron en variables cuantitativas, como se indica en la Tabla 1, con miras a identificar grupos de estudiantes con características comunes a partir de la observación de dichas variables en cada uno de ellos, y a su vez entender qué diferencia a los estudiantes de distintos grupos.

Nuestro diseño es de naturaleza no experimental y descriptivo. En el que se aplicó el método de encuesta y la técnica del cuestionario.

Para establecer los diferentes grupos de estudiantes, que darán origen a las diferentes tendencias de pensamiento probabilístico, hemos aplicado el análisis de clusters, mediante el método jerárquico de Ward y posteriormente aplicamos el método de partición de las k-Medias.

En segundo lugar, aplicamos el análisis discriminante descriptivo. Este se divide en dos partes que deben ordenarse en su aplicación. En primer lugar, determinamos si los tres grupos difieren significativamente entre sí considerando las variables de la Tabla 1. Este análisis es seguido por la obtención de las funciones discriminantes lineales y la evaluación de la importancia de cada variable en cada función discriminante a través de la matriz de estructura.

Tabla 1. Variables cuantitativas consideradas en el estudio.

A ₁₁	Cantidad de sucesos reconocidos correctamente como aleatorios y argumentados desde la categoría Causalidad por el estudiante.
A ₁₂	Cantidad de sucesos reconocidos correctamente como aleatorios y argumentados desde la categoría Multiplicidad por el estudiante.
A ₁₃	Cantidad de sucesos reconocidos correctamente como aleatorios y argumentados desde la categoría Incertidumbre por el estudiante.
A ₁₄	Cantidad de sucesos reconocidos correctamente como aleatorios y argumentados desde la categoría Subjetiva por el estudiante.
A ₂₁	Cantidad de sucesos reconocidos incorrectamente como no aleatorios desde la categoría Causalidad por el estudiante.
A ₂₂	Cantidad de sucesos reconocidos incorrectamente como no aleatorios desde la categoría Multiplicidad por el estudiante.
A ₂₃	Cantidad de sucesos reconocidos incorrectamente como no aleatorios desde la categoría Incertidumbre por el estudiante.
A ₂₄	Cantidad de sucesos reconocidos incorrectamente como no aleatorios desde la categoría Subjetiva por el estudiante.
P ₅	Cantidad de sucesos argumentados desde la Contingencia en la estimación de la probabilidad por el estudiante.
P ₆	Cantidad de sucesos argumentados desde la categoría Laplaciana en la estimación de la probabilidad por el estudiante.
P ₇	Cantidad de sucesos argumentados desde la categoría Frecuencial en la estimación de la probabilidad por el estudiante.
P ₈	Cantidad de sucesos argumentados desde la categoría Equiprobabilidad en la estimación de la probabilidad por el estudiante.
P ₉	Cantidad de sucesos argumentados desde la categoría Experiencial en la estimación de la probabilidad por el estudiante.

A continuación aplicamos el análisis discriminante predictivo; mediante la obtención de los coeficientes de clasificación lineal y luego con la evaluación de la clasificación.

III. RESULTADOS

A. Análisis de Clusters

Las variables presentadas en la Tabla 1 fueron medidas en los 81 estudiantes y usadas en la aplicación del análisis de clusters. El método aplicado fue el método jerárquico de Ward. Este método es apropiado para la constitución de grupos homogéneos respecto de las variables consideradas. Además, permite la construcción del dendograma, para la determinación del número de clusters. En nuestro caso obtuvimos tres clusters, como se muestra en la Fig. 1. A partir de la consideración de la presencia de 3 clusters se aplicó el método de las K-Medias. La diferencia fundamental de este método con un método jerárquico reside en que no supone que la asignación de un sujeto a un cluster sea irrevocable, un sujeto puede ser asignado a un cluster en una iteración y reasignado a otro cluster diferente en otro.

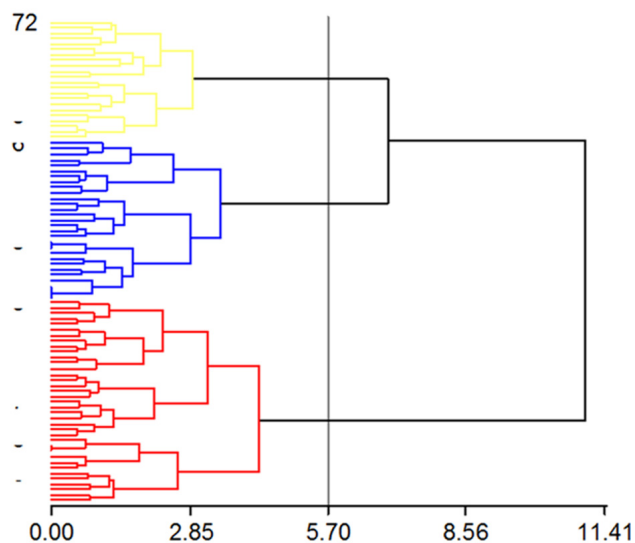


Fig. 1. Dendograma.

B. Análisis Discriminante

1) Test de igualdad de las medias de los grupos

Luego de determinar la pertenencia de cada estudiante al cluster correspondiente; aplicamos el análisis discriminante. Comenzamos realizando las pruebas de igualdad de las medias de las variables en los clusters, dando como resultado

que las medias de las variables A14, A21, P5, P9 no arrojaron diferencias significativas; por lo tanto, nuestro estudio se basará fundamentalmente en las variables que arrojaron diferencias significativas; es decir en, A11, A12, A13, A22, A23, A24, P6, P7 y P8.

2) Funciones discriminantes canónicas

Las funciones discriminantes son una combinación lineal de las variables predictoras definidas en la Tabla 1. Cuando se tienen tres grupos y más de dos variables predictoras, es posible definir dos funciones discriminantes. El objetivo es maximizar los valores propios de la matriz W-1B, o sea resolver la ecuación:

$$(W-1B-\lambda I)b=0 \quad (1)$$

donde W es la matriz de sumas de cuadrados y productos cruzados intragrupo o error; B es la matriz que contiene las sumas de cuadrados y productos cruzados del tratamiento, y b es el vector de pesos que maximiza la separación de los grupos. Así, obtuvimos los valores propios (Tabla 2) y los pesos de las funciones discriminantes estandarizadas:

Tabla 2. Valores propios.

Función	Valores propios	Porcentaje de varianza	Porcentaje de varianza acumulada	Correlación canónica
1	1,950	64,9	64,9	,813
2	1,055	35,1	100,0	,717

Tabla 3. WILKS' LAMBDA.

Test de funciones	Wilks' Lambda	Chi-cuadrado	Grados de libertad	Nivel de significación
1 - 2	,165	137,866	8	,000
2	,487	55,116	3	,000

Luego, las funciones discriminantes estandarizadas que se obtuvieron:

$$f1= 0.735A11+0.347A12-0.277A13+0.785P6$$

$$f2=0.064A11+0.496A12+0.969A13+0.282P6$$

La primera es la función que tiene mayor poder discriminante, explica el 64,9% de la variabilidad

del modelo, mientras que la segunda sólo explica el 35,1%, aunque según los valores p de la lambda de Wilks son significativas las dos funciones discriminantes (Tabla 3). Los coeficientes de estas funciones permiten comparar las variables. Las variables con coeficientes grandes en valor absoluto indican que tienen un alto poder discriminante. Además, se ha incluido un gráfico de dispersión para valorar cómo las dos funciones discriminantes conjuntamente contribuyen a separar los grupos.

3) Matriz de estructura

Tabla 4. Matriz de estructura.

Función	1	2
P ₆	,562*	,318
A ₁₁	,502*	-,111
A ₁₂	,277*	,192
P ₅ a	-,226*	,023
P ₈ a	-,120*	,000
A ₁₃	-,339	,848*
P ₉ a	-,223	-,317*
A ₂₂ a	,171	-,248*
A ₂₄ a	-,150	-,193*
A ₂₃ a	-,150	-,193*
P ₇ a	-,067	-,116*
A ₂₁ a	,012	-,044*
A ₁₄ a	,012	-,044*

Esta matriz (Tabla 4) revela qué variables son útiles para discriminar unos grupos de otros. Los elementos de la matriz son correlaciones intragrupos combinadas entre las variables discriminantes y las funciones discriminantes canónicas estandarizadas.

Las variables se han ordenado de acuerdo al valor de la correlación con la primera función discriminantes. Las variables A14, A21, A22, A23, A24, P5, P7, P8 y P9 no se han considerado en el análisis. El * indica la mayor correlación absoluta entre cada variable y cualquier función discriminante.

Para determinar las características de cada uno de los grupos, hemos calculado los valores medios de las variables en cada uno de ellos (Tabla 5). La denominación de las tendencias de pensamiento probabilístico está determinada por las categorías dominantes en cada grupo. El grupo 1 representa la tendencia al Determinismo, el grupo

2 la tendencia a la Causalidad y el grupo 3 la tendencia a la Incertidumbre.

Tabla 5. VALORES MEDIOS DE LAS VARIABLES.

Variable	Grupo 1	Grupo 2	Grupo 3
A ₁₁	2.89	0.83	0.82
A ₁₂	2.63	1.38	2.45
A ₁₃	2.07	2.63	6.18
A ₁₄	0.17	0.38	0.00
A ₂₁	0.17	0.38	0.00
A ₂₂	0.33	0.79	0.18
A ₂₃	0.46	0.88	0.09
A ₂₄	0.46	0.88	0.09
P ₅	3.37	3.96	3.73
P ₆	5.15	2.25	4.45
P ₇	2.04	2.83	1.91
P ₈	0.70	1.92	1.55
P ₉	0.59	0.83	0.36

4) *Mapa territorial*

El mapa territorial es un procedimiento para la representación gráfica de la clasificación de los estudiantes. Consiste en situar en el eje horizontal y en el vertical dos funciones discriminantes y separar por líneas las zonas que ocuparían los estudiantes clasificados en cada grupo. Cuando los estudiantes están bien clasificados, su representación sobre el plano los situaría en la zona correspondiente al grupo. En la Fig. 2, mostramos el mapa territorial de nuestra clasificación de los estudiantes en tres grupos.

5) *Análisis de resultados*

Los resultados de este estudio han revelado la presencia de tres tendencias de pensamiento bien definidas: tendencia hacia la Causalidad, tendencia hacia la Incertidumbre y tendencia hacia el Determinismo. El grupo de estudiantes que

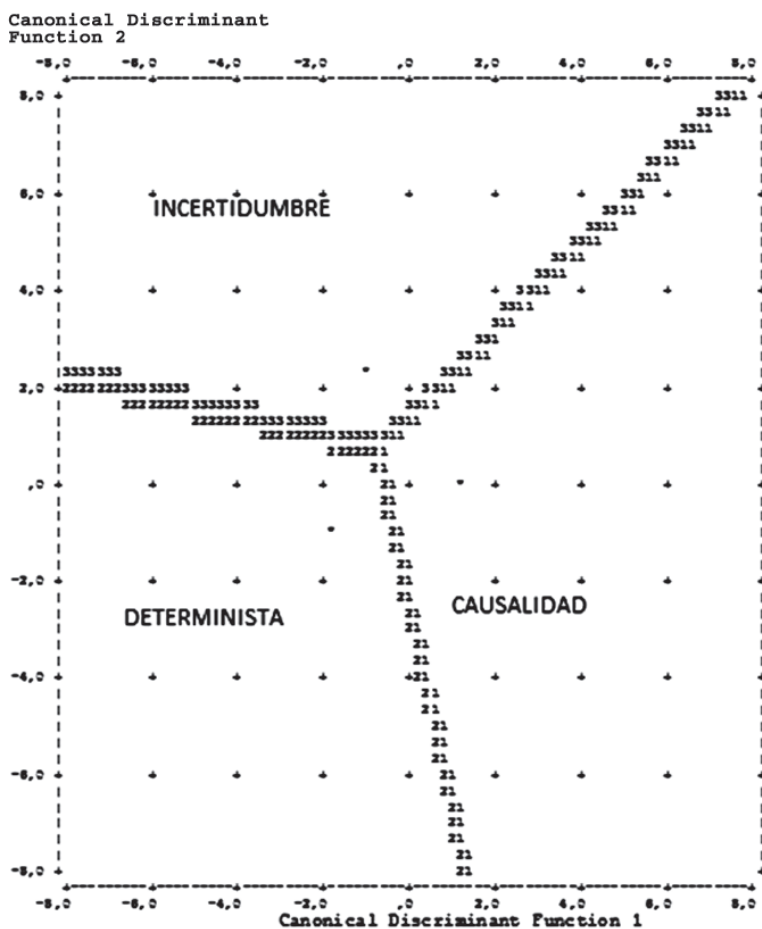


Fig. 2. Mapa territorial.

representa la tendencia al Determinismo, es el que logra el menor reconocimiento de la aleatoriedad, en promedio, un estudiante de este grupo reconoce correctamente sólo cinco sucesos como aleatorios de los doce propuestos. En relación con la estimación de la probabilidad, la categoría Frecuencial (P7) y la Equiprobabilidad (P8) alcanzan valores máximos en este grupo. Por lo que podemos concluir que basan la estimación de la probabilidad sólo en la lectura frecuencial del fenómeno o bien consideran que todos los posibles resultados tiene la misma probabilidad la ocurrencia del suceso. Está constituido por el 29,6% (24) de los estudiantes.

El grupo de estudiantes que representa la tendencia a la Causalidad, es el segundo grupo en el reconocimiento de la aleatoriedad, en promedio, un estudiante de este grupo reconoce aproximadamente ocho sucesos como aleatorios de los doce propuestos. Se destaca por argumentar la aleatoriedad desde la Causalidad (A11) y en segundo lugar desde la Multiplicidad (A12). Respecto de la estimación de la probabilidad, la categoría más usada es la Laplaciana (P6), la que alcanza el valor máximo en este grupo; mientras que la Equiprobabilidad (P8) alcanza el valor mínimo en este grupo. Es el grupo mayoritario que aglutina al 56,8% (46) de los estudiantes.

El grupo de estudiantes que representa la tendencia a la Incertidumbre, es el que logra el mayor reconocimiento de la aleatoriedad, en promedio, un estudiante de este grupo reconoce aproximadamente nueve sucesos como aleatorios de los 12 propuestos. Argumentan este reconocimiento de la aleatoriedad desde la Incertidumbre (A13); mientras que las categorías argumentativas de la negación incorrecta de la aleatoriedad alcanzan valores mínimos en este grupo. Cuando estima la probabilidad lo hace desde la categoría Laplaciana (P6) con un valor que supera a la media de esta variable, siendo inferior al alcanzado por el grupo Causalidad. Este grupo aglutina solamente al 13,6% (11) de los estudiantes.

IV. CONCLUSIÓN Y DISCUSIÓN

El entramado de las disciplinas que integran el diseño curricular del plan de estudio para la obtención del título de ingeniero en sistemas de in-

formación genera en los estudiantes del tercer año, tres modelos de pensamiento probabilístico. La tendencia al Determinismo integrada por casi la tercera parte de los estudiantes, la que representa una situación preocupante, debido a que estos estudiantes ya deberían ser consciente de la presencia de la incertidumbre en el mundo que los rodea, no solamente como futuros profesionales, sino también como ciudadanos de una sociedad impregnada por los fenómenos inciertos. En cuanto a la estimación de la probabilidad, se nota la presencia del sesgo de equiprobabilidad; a pesar de que los sucesos presentados, en la mayoría de los casos no corresponden a espacios muestrales equiprobables. Por lo que concluimos que no han logrado transferir la teoría normativa de la probabilidad a los fenómenos que surgen en la vida cotidiana y en el mundo físico-natural.

En relación con la tendencia hacia la Causalidad, si bien los estudiantes logran un nivel aceptable en el reconocimiento de la aleatoriedad, sus argumentos no son convincentes, buscan relaciones de causa-efecto para enjuiciar situaciones de azar, que surgen ante la necesidad humana de reducir o suprimir la incertidumbre en el entorno [5]. Coherentemente con el pensamiento causal que manifiestan, cuando estiman la probabilidad lo hacen desde la categoría Laplaciana. Recordemos que para Laplace todo fenómeno que ocurre tiene una causa determinada. Con Laplace se inicia un estilo de pensamiento, que evidentemente acompaña a la humanidad hasta nuestros días. El hombre racional de principios del siglo XIX pensaba que el mundo parecía azaroso sólo porque se desconocían todos sus mecanismos internos; quedaba así definido algo que podríamos llamar «azar subjetivo», ya que el observador es el origen de las incertidumbres. En este contexto, el concepto de probabilidad contribuía a suplir este desconocimiento. Este tipo de pensamiento está relacionado con la doctrina, que supone un universo determinista pero que aparece como aleatorio por la existencia de razones o causas ocultas, y que fue llamada «la doctrina de la necesidad» por S.C. Pierce [6], defensor de un azar objetivo.

Finalmente, la tendencia a la Incertidumbre, es la que logra el mayor reconocimiento de la aleatoriedad, pero es la que tiene menor presencia, si bien se esperaba que fuera la mayoritaria. Cuando estiman la probabilidad lo hacen desde la cate-

goría Laplaciana, con un uso menor al del grupo Causalidad. En esta tendencia el azar es intrínseco al mismo y considera que el mundo que nos rodea es irreductiblemente estocástico. En este caso, el azar reside fuera del sujeto que describe al mundo y por consiguiente lo podemos llamar «azar objetivo».

Desde nuestra perspectiva docente, consideramos que la enseñanza tardía de la probabilidad conjuntamente con un sistema educativo que entrena mayormente en el desarrollo de un pensamiento causal, son los motivos fundamentales de los resultados obtenidos en el presente trabajo. Sin embargo, los estudiantes deben reconocer que la Teoría de la Probabilidad es un marco teórico para la construcción de modelos matemáticos que tienen mayor poder explicativo que los modelos deterministas, en el estudio de los fenómenos que ocurren en el mundo real.

REFERENCIAS

- [1] A. Moreno (2014). El pensamiento Probabilístico en los Profesores de Biología en Formación. *Bolema: Boletim de Educação Matemática*.
- [2] A. Moreno. La Aleatoriedad en los Profesores de Biología y de Matemática en Formación: Análisis y Contraste de Significados. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*.
- [3] A. Moreno (2014). «Overview of Prospective Mathematics Teachers' Probabilistic Thinking». En 2014 K. Makar, B. De Sousa & R. Gould (Eds.), *ICOTS-9 Conference Proceedings. Sustainability in statistics education. 9th International Conference on Teaching Statistics*.
- [4] J.M Cardeñoso, «Las creencias y conocimientos de los profesores de primaria andaluces sobre la Matemática escolar. Modelización de conceptos sobre la aleatoriedad y probabilidad». Tesis doctoral. Universidad de Cádiz.
- [5] T.R. Krynski, Tenenbaum, J.B.(2007). «The Role of Causality in Judgment Under Uncertainty». *Journal of Experimental Psychology General*. 136(3), pp. 430-450.
- [6] I. Hacking, «La domesticación del azar: la erosión del determinismo y el nacimiento de las ciencias del caos». España: Gedisa editorial, 1991.

