

SIGNIFICADOS DE LA DISTRIBUCIÓN NORMAL EN LA UNIVERSIDAD

Miriam Carpio*, Cecilia Gaita*
Miguel R. Wilhelmi, Álvaro Sáenz de Cabezón****

**Pontificia Universidad Católica del Perú*

***Universidad Pública de Navarra*

Resumen

El principal indicador empírico de las dificultades de aprendizaje en los procesos de instrucción es la desadaptación entre los significados personales mostrados y los significados institucionales evaluados. Analizamos esta desadaptación con relación a los significados de la distribución normal en un contexto universitario. Para ello caracterizamos el significado de referencia, que enmarca los significados institucionales pretendidos y efectivamente implementados. Dado el carácter eminentemente formalizado del significado institucional pretendido, la valoración de la complejidad ontosemiótica del objeto “distribución normal” se ve condicionada por prácticas operatorias previas. Las restricciones ecológicas se revelan entonces esenciales para el análisis del acoplamiento entre los significados personales y el significado institucional.

Palabras clave: distribución normal, significados institucionales, dualidad personal-institucional, complejidad ontosemiótica.

Abstract

The main empirical indicator of the difficulties of learning in the instruction processes is the mismatch between the evaluated institutional meaning personal shown and the meaning. We analyze this mismatch in relation to the meaning of the normal distribution in a university context. For it, we characterized the reference meaning, that frames the meaning institutional tried and indeed implemented. Given the character especially formalized of the tried institutional meaning, the valuation of the onto-semiotic complexity of the object “normal distribution” is conditional by previous operating practices. The ecological restrictions reveal then essential for the analysis of the connection between the meaning personal and the institutional meaning.

Key words: normal distribution, institutional meaning, personnel-institutional duality, onto-semiotic complexity.

Introducción

La inferencia estadística está fundamentada sobre el estudio de las distribuciones de probabilidad donde aparecen conceptos tan abstractos como los de función de densidad y de distribución (Tauber, 2001). En este contexto, la distribución normal tiene un papel central, ya que, por el teorema central del límite, se puede asegurar que es una buena aproximación de otras distribuciones de probabilidad para muestras de tamaño suficientemente grande. Nuestro objetivo es la identificación de los factores que condicionan la enseñanza y el aprendizaje de la distribución normal en la universidad.

Dada la complejidad matemática de la noción “distribución normal”, la inferencia estadística necesita de conocimientos previos profundos tanto de análisis matemático, como de cálculo de probabilidades para poder generalizar los resultados de una muestra a toda una población. Estos requerimientos pueden no haberse cumplido al iniciarse un curso introductorio de estadística.

Para el análisis del acoplamiento progresivo entre los significados personales y el institucional se determina, en primer lugar, el significado de referencia a partir del estudio epistémico de la distribución normal en un libro de texto y se realiza una mención especial de los programas informáticos; luego se define el significado pretendido teniendo como base una muestra de textos de probabilidad y estadística inferencial. Para la descripción del significado institucional implementado, se observa la secuencia de enseñanza. Las herramientas de evaluación empleadas en la secuencia didáctica permiten describir el significado evaluado, que es contrastado con los significados personales declarados por los estudiantes.

Contexto

Según Tauber (2001), en la actualidad el primer contacto con la estadística debería darse a través de datos reales e incluso construyendo un sistema de datos propio para su análisis. En particular, en un curso universitario de introducción a la estadística deberían plantearse situaciones- problema en las que sea necesario realizar, con ayuda de un paquete estadístico, un análisis exploratorio de los datos que permita posteriormente interpretar resultados.

Esta actividad contrasta radicalmente con la actividad matemática en las universidades peruanas, basada casi exclusivamente en la resolución de problemas de cálculo rutinarios presentados en los libros de texto. Una prueba de ello se encuentra en la experiencia analizada: un curso introductorio de estadística y probabilidad en la Universidad de Piura, con sede en Lima. Los 45 estudiantes provenían de las especialidades de economía y administración.

Objetivos

El objetivo general es la identificación de los factores que condicionan la enseñanza y aprendizaje de la distribución normal en el nivel superior, considerando que la enseñanza implica la participación del estudiante en las comunidades de prácticas que soportan los significados institucionales y que el aprendizaje supone la apropiación del estudiante de dichos significados. Esto, bajo el supuesto que debe haber un acoplamiento progresivo entre los significados personales e institucionales.

Los objetivos para esta investigación son los siguientes:

- Caracterizar el significado de referencia para la distribución normal.
- Analizar la secuencia de enseñanza propuesta por el docente y los libros de texto empleados en el curso con el objeto de determinar los significados pretendidos en el proceso de estudio.

- Describir los elementos de significado efectivamente observados en el desarrollo de la secuencia de enseñanza para determinar el significado institucional implementado.
- Analizar los comportamientos de los estudiantes mostrados en la evaluación al finalizar el proceso de instrucción de la distribución normal para contrastarlos con los significados institucionales originariamente pretendidos.

Marco teórico

En el Enfoque Ontosemiótico sobre el conocimiento y la instrucción matemáticos (EOS), (Godino, Batanero y Font, 2008) la definición de una ontología de los objetos matemáticos precede a la definición de herramientas teóricas para el análisis de los sistemas didácticos. Los aspectos y cuestiones básicas planteadas en la teoría son (Montiel, Wilhelmi, Vidakovic y Elstak, 2009):

1. El objeto matemático (*¿qué es un objeto matemático?*) y la definición de una ontología matemática (*¿qué entidades matemáticas deben ser consideradas?*). Un objeto es cualquier entidad utilizada, explícita o implícitamente, en una práctica matemática. En el EOS se definen seis entidades primarias: lenguaje, situaciones, propiedades, proposiciones, acciones y argumentaciones.
2. La ecología de los objetos matemáticos (*¿cuál es la realidad contextual y funcional de los objetos matemáticos?*). Los objetos matemáticos son analizados según cinco facetas duales: personal / institucional; ostensivo / no-ostensivo; ejemplar / tipo; unitario / sistémico; expresión / contenido.
3. La relación entre los objetos y la determinación de configuraciones (*¿cómo se estructuran los objetos matemáticos en las prácticas matemáticas?*). Los

objetos constituyen configuraciones epistémicas o cognitivas, según la realidad institucional o personal de la práctica matemática.

4. Tipos de significados según los agentes (*¿cuál es la realidad —personal o institucional— de los objetos matemáticos?*). Para el significado de los objetos matemáticos se consideran los *sistemas de prácticas* (operativas y discursivas) que manifiestan las personas para resolver tipos de situaciones problemáticas; este sistema de prácticas puede ser realizada por una persona o compartidas por el seno de una institución, lo que determina los significados personales e institucionales.
5. Emergencia de objetos en la actividad matemática (*¿cómo el conocimiento matemático es generado a partir de la relación entre objetos?*). La emergencia se vincula con la relación entre configuraciones mediante un mecanismo de significación (expresión-contenido): una configuración epistémica (previa) lleva a otra (emergente) y esta emergencia se “muestra” en la movilización de un conocimiento en la realización de una tarea o la resolución de un problema.

La comprensión matemática es considerada como una competencia: un sujeto comprende un objeto matemático cuando lo usa de manera competente en las prácticas que no suelen ser las habituales a las típicas.

El EOS considera como *práctica matemática* a toda actuación o expresión realizada por alguien para resolver problemas matemáticos, comunicar a otros la solución obtenida, validarla o generalizarla a otros contextos y problemas (Godino y Batanero, 1994, 334).

En los significados institucionales se proponen los siguientes tipos:

- *Referencial*: significado de referencia del significado pretendido.
- *Pretendido*: sistema planificado incluido en el proceso de estudio.

- *Implementado*: proceso de estudio implementado por el docente.
- *Evaluado*: subsistema que utiliza el docente para evaluar el aprendizaje.

En cuanto a los significados personales, se consideran los siguientes:

- *Global*: totalidad del sistema de prácticas personales que se puede manifestar ante un objeto matemático.
- *Logrado*: prácticas manifestadas y que están conforme las establecidas institucionalmente.
- *Declarado*: prácticas efectivamente expresadas en las evaluaciones.



Fig.1: Tipos de significados institucionales y personales

En la figura 1 (Godino, 2008, p.6), se puede ver la tipología básica de significados y cómo se relacionan éstos en un proceso de enseñanza y aprendizaje donde debe presentarse el acoplamiento progresivo entre los significados personales e institucionales.

Método

La primera etapa de la investigación es descriptiva y exploratoria, siendo el foco de atención los elementos contextuales y procesos cognitivos/epistémicos. Basándonos en Alvarado (2007), se muestran elementos específicos (tales como la determinación de campos de problemas y la organización de términos y

expresiones verbales) para el análisis sobre el significado y comprensión de la distribución normal.

En la segunda etapa, se emplean algunos elementos cuantitativos para organizar las respuestas del cuestionario de evaluación. Luego, siguiendo un enfoque interpretativo y explicativo, se contrastan los comportamientos esperados y los errores previsibles con los realmente observados, basándonos en la *ingeniería didáctica* (Artigue, 1989) como método de investigación.

Significados de referencia de la distribución normal

Los campos de problemas identificados muestran un abanico extenso de aplicaciones de la normal. Sin embargo, en la mayoría de los casos los problemas aparecen descontextualizados, habiendo desaparecido el origen de los mismos. ¿Por qué es necesaria la aproximación a la normal de distribuciones discretas y el cálculo por estandarización de la normal? Dos aspectos de distinta naturaleza:

1. *Técnica*, imposibilidad de tabular todas las funciones normales.
2. *Teórica*, la importancia de relacionar distintas distribuciones de probabilidad y comparar poblaciones normalmente distribuidas.

Sin embargo, ¿es necesario “reducir” o “aproximar” a la normal estándar supuesta la existencia de programas informáticos?

Las nociones de experimento aleatorio y variable aleatoria son consustanciales a las distribuciones de probabilidad, en particular, de la normal. Los programas informáticos permiten poner de manifiesto estos aspectos por la generación de muestras aleatorias para una distribución de probabilidad y observando regularidades en las mismas.

Significados pretendidos e implementados de la distribución normal

Existe relación entre los significados pretendidos y los implementados, aunque debe considerarse que no todos los elementos de significados se trabajaron con la misma intensidad o en el orden en que fueron planificados. Esto se debe básicamente a que los alumnos no manejan algunos conceptos previos relacionados con la estadística y con el cálculo.

Se presentan definiciones de la distribución normal, de la distribución normal estándar y de la distribución acumulada. Para ello hay que tener en cuenta los elementos de significado:

- *Campos de problemas*
 - CP1: Poblaciones con distribución normal
 - CP2: Aproximación de distribuciones discretas o continuas a la distribución normal. Aproximación de la binomial a la normal.
 - CP3: Obtención de la distribución en el muestreo de la media

En el campo CP2 sólo se considera la aproximación de la binomial a la normal como un subcampo; no se aproxima la distribución hipergeométrica, la poisson o la chi-cuadrado a la normal.

En cuanto al CP3, este se aborda directamente empleando la fórmula correspondiente.

- *Lenguaje*

La simbología asociada a los parámetros y estadísticos, a la función de densidad y al cálculo de probabilidades es densa. La propiedad reproductiva de la normal, que requiere el trabajo con varias variables, se presenta haciendo uso de una simbología también densa.

En la mayoría de casos se emplean las representaciones gráficas para la comprobación o generalización de resultados. Ocasionalmente se presentan algunas demostraciones deductivas formales.

- *Conceptos*

Se introducen nuevos términos tales como variable aleatoria continua, distribución de probabilidad, parámetros, media, desviación estándar, varianza, etc., cuya comprensión será necesaria para la de la distribución normal.

- *Propiedades*

La propiedad simétrica de las probabilidades se representa gráficamente y se emplea en la solución de algunos campos de problemas. Solo en la solución de algunos campos de problemas se recurre a la gráfica de la distribución normal.

- *Procedimientos*

Al introducir los procedimientos en la resolución de problemas como la estandarización, el uso de la tabla normal, el cálculo de probabilidades y el cálculo de los valores Z , se observa el predominio de prácticas algorítmicas, acompañadas de cálculos algebraicos.

Cuestionario

1. Suponga que el ingreso mensual por familia en miles de soles, es una variable aleatoria X con función densidad de probabilidad:

$$f(x) = \begin{cases} 0.25x & \text{si } 0 \leq x \leq 2 \\ 1 - 0.25x & \text{si } 2 < x \leq 4 \end{cases}$$

Considerando una muestra aleatoria de 100 ingresos:

- a) Calcule la probabilidad de que el ingreso total de la muestra supere los 216 mil soles.

- b) Halle la probabilidad de que el ingreso promedio de la muestra sea menor que 3 mil soles.
2. El precio de venta de un artículo es una variable aleatoria X , que tiene una media de \$10 y una desviación estándar de \$2. Si se adquieren 64 artículos para su comercialización, halle el precio de compra K de cada uno de ellos de manera que al vender todos, se quiere ganar al menos \$192 con probabilidad de 0.975.
3. La utilidad en miles de soles por la venta de cierto artículo, es una variable aleatoria con distribución normal. Se estima que en el 5% de las ventas la utilidad sería menor que 6.71 mil soles, mientras que en el 1% de las ventas la utilidad sería mayor que 14.66 mil soles. Si se realizan 16 operaciones, calcule la probabilidad de que la utilidad promedio esté entre 10 y 11 mil soles.
4. El índice de cotización diaria de cierta acción en la bolsa de valores de Tokio, sigue una distribución normal con media de 500 soles y desviación estándar de 25 soles. Un accionista invertirá cuando el índice de cotización diaria se sitúa entre 475 y 550 soles. Halle la probabilidad de que invierta en un día.

Nos referimos al primer ejercicio únicamente. En la tabla 1 se pueden ver los comportamientos esperados y errores previsibles de los estudiantes, clasificados según la entidad primaria a la que hacen referencia.

Tabla 1. Comportamientos esperados

Comportamiento	Entidad
A: La utilización de las propiedades de probabilidad son correctas en el apartado a), teniendo en cuenta que el planteamiento puede ser correcto o no conforme a lo que pide el problema.	Propiedad
B: La variable aleatoria inicial del problema es considerada en el planteamiento de una forma correcta, según lo enseñado en el proceso de estudio.	Concepto-definición
C: Reconoce la nueva variable aleatoria Y del apartado a). Puede que esta definición sea explícita.	Concepto-definición
D: Plantea en forma correcta las propiedades y definiciones que permiten hallar la media y varianza. Independientemente si los cálculos realizados son o no correctos.	Propiedad
E: Dispone de los valores correctos para la media y la varianza (indistintamente de cuál es el origen de estos valores: cálculo correcto, uso de calculadora, o copia).	Procedimiento
F: Asocia el cálculo de la media y la varianza con una integral y la resuelve de manera correcta.	Procedimiento
G: Reconoce la nueva variable aleatoria $\bar{Y}$ del apartado b). Puede que esta definición no sea explícita.	Concepto-definición
H: La función de densidad es reconocida tal como se define en el problema. No hay relación con otras distribuciones (tales como la normal, binomial, etc.) que no correspondan a la función de densidad definida por el problema.	Situación
I: Se llega a la respuesta esperada del apartado a). No se considera el procedimiento seguido.	Procedimiento
J: Se llega a la respuesta esperada del apartado b). No se considera el procedimiento seguido.	Procedimiento
K: En el proceso de estandarización del apartado a), el alumno no confunde la varianza por la desviación estándar.	Concepto-definición
L: Se define correctamente la varianza del apartado b).	Concepto-definición
M: En la estandarización del apartado b), el alumno no confunde la varianza por la desviación estándar.	Concepto-definición

Significados personales en la evaluación final

Resultados

En la tabla 2 se puede ver la frecuencia de los comportamientos en la prueba final aplicada a la muestra objeto de estudio. Se indica:

- *Fa*: frecuencia absoluta de alumnos manifiestan un comportamiento.

- *Blanco*: frecuencia absoluta de alumnos que no presentan un comportamiento.
- *Error*: frecuencia absoluta de alumnos que comenten un error relacionado con un determinado comportamiento.

Tabla 2. Frecuencias absolutas de los comportamientos

Comportamiento	Fa	Blanco	Error
A	22	7	2
B	21	4	6
C	20	7	4
D	20	5	6
E	19	6	6
F	19	5	7
G	18	11	2
H	17	5	9
I	15	8	8
J	14	11	6
K	12	8	11
L	11	12	8
M	9	12	10

Los procedimientos algebraicos y la estandarización fueron, en general, los mejores utilizados en la solución de los problemas por parte de los estudiantes.

Por otro lado, se detectó que varios elementos de significados no fueron utilizados correctamente por los estudiantes. Entre los más frecuentes se encontró la propiedad reproductiva de la normal, donde se presentó un conflicto entre la varianza y la desviación estándar. Algo similar ocurrió al aplicar el teorema del límite central. Estos errores pueden tener su origen en los errores cometidos por el docente en la formulación de algunos ejemplos y situaciones.

En general, el análisis de las diferencias entre los distintos significados institucionales permitió prever algunas dificultades y conflictos semióticos de los estudiantes. Así como, el análisis de los significados personales proporcionó

información sobre la forma correcta o incorrecta de algunos elementos de significado que usaron los estudiantes.

Discusión de los resultados

La función de densidad asociada a la pregunta 1, no representa una distribución normal. Se esperaba que los alumnos, haciendo uso de la representación gráfica, llegaran a este resultado. Sin embargo, se fijaron en el tamaño de la muestra involucrada en el problema. Esto se debe a que la representación gráfica ha sido utilizada únicamente para la introducción de la definición de la distribución normal y no ha cumplido ningún otro papel.

Uno de los comportamientos esperados era (*a priori*) que los alumnos tuvieran dificultad en la definición de los parámetros asociados a las nuevas variables. Sin embargo, la definición de las nuevas variables fue también problemática (35% de fracaso). Este hecho permite afirmar (al menos para la muestra de estudiantes seleccionada) que la definición de variables a través de la sumatoria es conflictiva. En particular, el promedio muestral de “n” involucra un nuevo significado relacionado con la teoría de las distribuciones de muestreo.

Por otro lado, dado que el curso analizado no ha sido diseñado intencionalmente, se justifica que el significado pretendido difiera del de referencia y del implementado.

Así, mientras que para el significado de referencia se plantea el uso de un software que permita abordar los problemas del CP1 (problemas con distribución normal) de una manera más económica, en el significado implementado se recurre a procesos de estandarización y al uso de las tablas de la distribución normal.

El optar por el uso de un software reduciría el tiempo empleado en procedimientos rutinarios. Esto permitiría introducir más preguntas relacionadas

con la comprensión de la distribución normal y de sus propiedades, así como preguntas donde se deban interpretar los resultados. Estos campos de problemas no han sido considerados en el proceso de estudio.

En esta misma dirección, es preciso brindar a la representación gráfica un papel operativo, discursivo y regulativo. Su uso restringido para la introducción de la definición de la distribución normal, no permite generar condiciones para la construcción y comunicación de conocimientos relativos a las distribuciones de probabilidad.

A modo de conclusión: cuestiones abiertas

La identificación de desadaptaciones entre los significados personales e institucionales pretendido, implementado y evaluado, así como entre los significados institucionales de referencia y pretendido de la distribución normal sugiere la necesidad de elaboración de un proceso instruccional potencial factible. Este proceso se basará en los aspectos de mejora detectados de la idoneidad didáctica, entre los que se encuentra el uso de los programas informáticos. Asimismo, la puesta en marcha del proceso permitirá extraer conocimientos adicionales de los procesos de construcción y comunicación de los significados asociados a las distribuciones de probabilidad, en particular de la distribución normal.

Reconocimiento: Trabajo realizado en el marco del proyecto de investigación, SEJ2007-60110/EDUC. MEC-FEDER.

Referencias

- Alvarado, H. (2007). Significados del Teorema Central del Limite en la Enseñanza de la Estadística en Ingeniería, Tesis doctoral. Granada: Universidad de Granada.
- Artigue M. (1989), Ingénierie didactique. *Recherches en Didactique des Mathématiques*, 9(3), 282–307.
- Córdova, M. (2006). *Estadística descriptiva e inferencial*. Lima: Editorial Moshera.
- Godino J. D., Batanero C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques* 14 (3): 325-355. [Institutional and personal meaning of mathematical objects. *Journal für Mathematik-didaktik*, 1996, 99-121].
- Godino, J. D., Batanero, C. y Font, V. (2008). Un enfoque ontosemiótico del conocimiento y la instrucción matemática. Departamento de Didáctica de la Matemática. Universidad de Granada. Disponible en Internet: URL: http://www.ugr.es/local/jgodino/indice_eos.htm
- Montiel, M.; Wilhelmi, M. R.; Vidakovic, D.; Elstak, I. (en revisión). Vectors, Change of Basis and Matrix Representations of a Linear Transformation: the Onto-Semiotic Approach in the Analysis of Creating Meaning. *Journal for Research in Mathematics Education*, 2009.
- Novales, A. (1996). *Estadística y econometría*. Madrid : McGraw-Hill.
- Spiegel, M. (1991). *Estadística*. Madrid : McGraw Hill.
- Tauber, L. (2001). *La construcción del significado de la distribución normal a partir de actividades de análisis de datos*, Tesis doctoral. Sevilla: Universidad de Sevilla.