

EL CONCEPTO DISTRIBUCION DE PROBABILIDAD: SU EVOLUCION EN ESTUDIANTES DE INGENIERIA



Sánchez C. Tomás, Albert J. Armando

tsanchez@itesm.mx, albert@itesm.mx

Instituto Tecnológico y de Estudios Superiores de Monterrey

Resumen

Un concepto fundamental de la probabilidad y la estadística es el de distribución de probabilidad. Existen varios estudios que tratan de clarificar el concepto y enfatizan sobre la necesidad de reconocer curricularmente su importancia, otros que reportan las dificultades de los estudiantes al intentar construir esa estructura conceptual, así como aquellas investigaciones que sugieren rutas para que los estudiantes logren el objetivo. En este estudio se exploran los argumentos que estudiantes de ingeniería expresan al ser cuestionados, de forma escrita y gráfica, sobre el significado del concepto de distribución con el propósito de identificar procesos cognitivos que se dan al monitorear la evolución de esta noción durante el desarrollo de un curso de probabilidad y estadística.

Palabras Claves

Distribución, conceptual, probabilidad, estadística.

Introducción

Las distribuciones de probabilidad de una variable aleatoria constituyen una parte importante de cualquier curso de probabilidad y estadística para estudiantes de ingeniería. En los libros de texto para dichos cursos con seguridad encontraremos el tema de las distribuciones de probabilidad, para variables aleatorias discretas y continuas, tratado el tema más o menos como el texto siguiente tomado de Devore (2004): “La distribución de probabilidad o función de probabilidad de masa (pmf: probability mass function) de una variable aleatoria discreta está definida para todo número x por $p(x)=P(X=x)=P(\text{toda } s \in S: X(s)=x)$ ” o bien, en el caso de las distribuciones continuas, en Montgomery (1996) se encuentra la siguiente definición: “Una

función $f_x(x)$ es una función de densidad de probabilidad de la variable aleatoria continua X si para cualquier intervalo de números reales $[x_1, x_2]$

$$(1) f_x(x) \geq 0$$

$$(2) \int_{-\infty}^{\infty} f_x(x) dx = 1$$

$$(3) P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f_x(u) du ,$$

Es común también encontraremos el tema de las distribuciones muestrales (distribuciones en el muestreo) en los libros de texto en uso para el área de ingeniería por ejemplo en la obra citada, Pág. 286 se establece primero que *una estadística es una variable aleatoria que depende de los valores obtenidos en cada muestra particular* y dado que una estadística es una variable aleatoria, entonces tiene una distribución de probabilidad y establece, Pág. 287, la siguiente definición: “La distribución de probabilidad de una estadística recibe el nombre de distribución de muestreo”. De lo anterior podemos considerar que las distribuciones en el muestreo son “una especie de aplicación” del concepto de distribución. Concepto que se reconoce como clave en la estadística más aún se vislumbra que la distribución provee una fuerte conexión entre la estadística y la probabilidad, conexión que actualmente no es tomada en cuenta dentro de los planes de estudio escolares ni tampoco en la enseñanza, como lo dicen Pfannkuch y Reading (2006) diciendo además que partiendo del hecho de que la distribución un concepto clave en estadística, pareciera ser que tanto estadístico como educadores no han notado lo difícil que es para los estudiantes desarrollar un entendimiento profundo, conceptual y operacional de esta estructura conceptual; “distribución de probabilidad”.

Aceptándose la importancia de las mismas se debería aceptar también la necesidad de estudios más amplios y profundos para desentrañar las dificultades en el proceso de formación de este concepto. No es mucho lo que se sabe acerca de cómo es que fue evolucionando en los estudiantes el concepto de distribución de probabilidad y de cómo es que se fue concretando esa estructura conceptual subyacente. Wild (2006) dice que se puede encontrar mucho de los

que se ha escrito sobre los usos especializados y las definiciones de “distribución” pero casi nada de la “distribución” en sí, como una estructura conceptual subyacente.

Otros investigadores han hecho esfuerzos por conceptualizar la noción de distribución, por ejemplo Wild (2006) desarrolla la noción de distribución como una lente a través de la cual los estadísticos ven la variación en los datos, agregando que las graficas y los resúmenes que comúnmente se usan son formas de ver, resumir y comunicar los aspectos de la información presente en dichas distribuciones y concluye, pág. 22, fundamentalmente la noción de “distribución” es el patrón de variación en una variable (o en un conjunto de variables.). Por su parte otros investigadores han intentado ubicar los factores que pudieran estar contribuyendo, positiva o negativamente, al desarrollo de este concepto, por ejemplo Watson (2009) realizó una investigación con estudiantes de 6 a 15 años buscando ver la influencia de la variación y el valor esperado en el desarrollo conceptual de la distribución, encontrando -contrario a lo que comúnmente se cree- que los estudiantes se familiarizan mas fácilmente con la variación que con el valor esperado, pero no hay un resultado claro acerca de cómo influyen en el desarrollo del concepto de distribución.

Esta investigación se centra en el estudio de las concepciones acerca de la noción de distribución y su evolución en estudiantes de ingeniería. Aunque se reconoce la complejidad sistémica e importancia de abordar el fenómeno didáctico desde sus distintas componentes: epistemológica, cognitiva, de enseñanza y de sus prácticas sociales, este estudio se centra en la componente cognitiva vinculada con el ámbito cultural. Particularmente, en esta investigación se instrumentó una categoría de análisis de las prácticas sociales: la argumentación.

Se observaron y analizaron los argumentos que dieron los estudiantes ante situaciones problemáticas intencionalmente seleccionadas, para tres momentos específicos durante el desarrollo de un curso de Probabilidad y Estadística. Con el propósito de tener antecedentes en cuanto a la concepción de los alumnos acerca del término distribución se realizó una evaluación previa como referente a la observación y análisis de los argumentos de los estudiantes dentro del curso. Se realizó también una evaluación con alumnos de semestres superiores con la

intención de tener información que nos permita tener elementos para prospectar la permanencia o no del nivel de conceptualización alcanzada durante el curso y, a fin de cuentas, con la pretensión de descubrir más elementos del proceso cognitivo que se da cuando los estudiantes responden reactivos relacionados con los conocimientos adquiridos en un curso de probabilidad y estadística, se pretende observar las argumentaciones que los alumnos ponen en juego en torno a la noción de distribución, se trata de diferenciar o al menos ver la presencia de las creencias cotidianas y las influenciadas por la clase.

Metodología

Este estudio se realizó con alumnos de ingeniería de una Universidad del Norte de México en la que todos los estudiantes de las diversas carreras de ingeniería (un total de 21) tienen en su plan de estudios un curso común de Probabilidad y Estadística. Para varias de las carreras éste será un curso requisito para otros como Control Estadístico de Calidad, Análisis y Diseño de Experimentos, Investigación de Operaciones, etc., para otras carreras éste será el único curso de esta área en su plan de estudios. En esta Universidad, como la mayoría en nuestro país, los estudiantes ingresan sin antecedente académicos en la componente aleatoria del currículo.

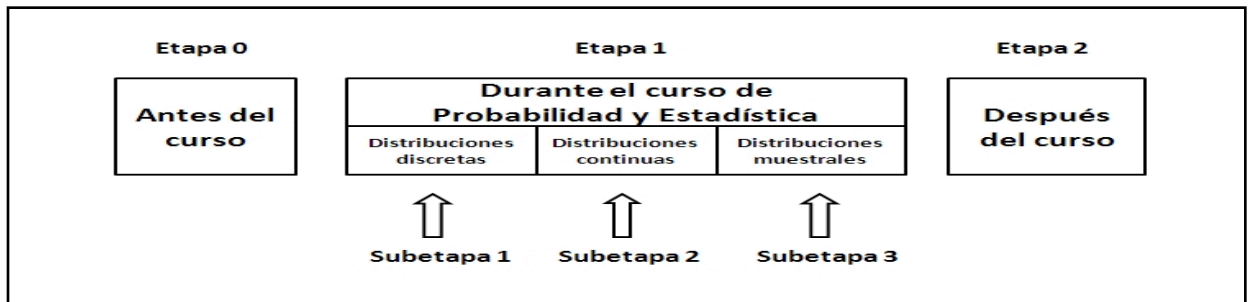
El estudio se realizó en varias etapas, las cuales son numeradas en función de los antecedentes académicos de esta temática.

- Etapa 0: Antes del curso de Probabilidad y Estadística
- Etapa 1 Durante el curso de Probabilidad y Estadística
- Etapa 2: Después de dos o más semestres del curso Probabilidad y Estadística

La Etapa 1 se subdividió en tres subetapas, correspondiendo cada una de ellas a momentos especiales de los contenidos del curso:

- Subetapa 1: Después de ver el tema de las distribuciones discretas.
- Subetapa 2: Después de ver el tema de las distribuciones continuas.
- Subetapa 3: Después ver el tema de las distribuciones muestrales.

El esquema completo se ilustra a continuación;



En la Etapa 0 participaron dos grupos con un total 68 estudiantes de nuevo ingreso a la Universidad, son alumnos sin antecedentes académicos en esta área y con la aplicación de las dos preguntas pretendíamos explorar sus creencias, concepciones que pensamos son las del dominio común (de acuerdo a las condiciones descritas). Las respuestas servirán como un referente para las etapas posteriores. Las preguntas son:

- Describe brevemente lo que significa para ti el concepto de distribución
- Como sería para ti una gráfica de una distribución (hacer un dibujo)

En la Etapa 1 se describe a continuación los participantes y el tipo de cuestionario por subetapas. En todas las subetapas participaron un total de cuatro grupos académicos. En la Subetapa 1 y 2 se busca explorar el cómo pudieran haber evolucionado las argumentaciones del dominio común a las opiniones influenciadas por la clase:

Subetapa 1: Se aplicó el mismo cuestionario de la Etapa 0 a un total de 116 alumnos de las diferentes carreras

Subetapa 2: Se aplicó el mismo cuestionario de la Etapa 0 a los mismos cuatro grupos participaron ahora 107 alumnos de las diferentes carreras de ingeniería.

Subetapa 3: En esta subetapa se aplicaron dos tipos de evaluación a los mismos cuatro grupos de las subetapas 1 y 2, se pretendía observar la evolución en el proceso de construcción de la noción de distribución, para ellos se indagó en el ámbito cualitativo (conceptual) y cuantitativo. La primera se aplicó a solo uno de los cuatro grupos, 29 alumnos, de las subetapas anteriores y consistió en tres preguntas abiertas:

Describe brevemente lo que significa para ti el concepto de distribución

Es llevar algo producido de un lugar a otro.

Y los que argumentaron que distribución tenía que ver con una acción de agrupar, clasificar o formar conjuntos, ejemplo:

Describe brevemente lo que significa para ti el concepto de distribución

= significa separar los conceptos de acuerdo a sus diferencias y cualidades

Los resultados de la parte conceptual muestran el dominio de conceptualización de la noción de distribución como una repartición, más específicamente se observa la presencia de una actividad, la de llevar, la de dar. La distribución incluye, de manera implícita y explícita, la acción de repartir. En el aspecto gráfico se encontró más variabilidad de respuestas, la mayoría girando alrededor de gráficos de pastel, diagramas de barras y coordenadas rectangulares.

Etapas 1

Subetapa 1. Se organizó la información, de las respuestas por escrito, según los argumentos dados por los alumnos en dos grandes grupos: Argumentos Cotidianos (un 44% aproximadamente) las semejantes a la Etapa 0 (repartir, entregar, etc.) y la Argumentos de la Clase (el 56% restante) aquellos que aludían a expresiones típicas del inicio de un curso de probabilidad y estadística; fundamentalmente los que hacían referencia al tema de las distribuciones discretas presentadas hasta ese momento en el curso. De manera similar el aspecto gráfico se organizó en Diagrama de la Clase, un 67% aproximadamente que fundamentalmente hacían referencia a diagramas de barras como el que se muestra, aunque no todos erróneos y Diagrama Cotidiano (17%) cuando hacían gráficos semejantes a los de la Etapa 0.

A continuación se dan dos ejemplos, el segundo es un típico caso de la Etapa 0 pero el primer ejemplo es una especie de transición entre la Etapa 0 y la Etapa 1 (entre la acción de repartir y las argumentaciones influidas por la clase). Destacamos el hecho de que en ambos ejemplos continua apareciendo como argumento la acción de repartir, el primero dice: “la repartición de

algún valor asignado entre un número determinado de objetos, valores, etc.”, se clasificó como Argumento de la Clase porque un renglón después dice: “la dist[ribución] de probabilidad 1 entre varios valores que puede tomar una variable”. El anterior es el caso de aquellos alumnos que usan su concepción antecedente y comienzan a incorporar en sus argumentaciones los proveídos por la clase. A lo anterior seguramente ayuda el hecho de que la estructura matemática de lo que es una distribución discreta no contradice el proceso o acción de repartir.

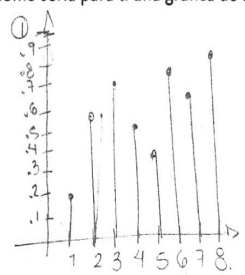
Nombre (Opcional): _____

Describe brevemente lo que significa para ti el concepto de distribución

Lo repartición de algún valor asignado entre un número determinado de objetos, valores, etc.

Lo dist de probabilidad 1 entre varios valores que puede tomar una variable

Como sería para ti una grafica de una distribución (hacer un dibujo)

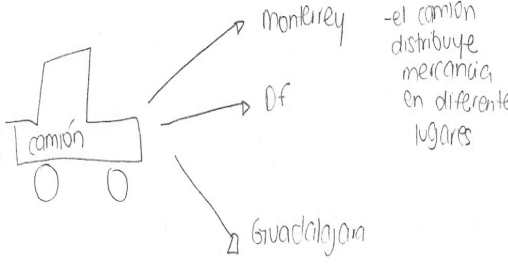


Nombre (Opcional): _____

Describe brevemente lo que significa para ti el concepto de distribución

Distribución significa repartir en diferentes áreas.

Como sería para ti una grafica de una distribución (hacer un dibujo)



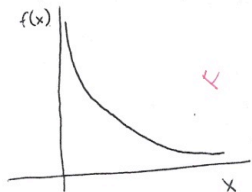
1. **Subetapa 2.** Las respuestas se agruparon de manera similar a la Subetapa 1, notando que ahora se amplía las posibilidades para los Argumentos de la Clase y el Diagrama Formal en vista de que hasta este punto ya se había cubierto el tema de las distribuciones continuas. En esta parte del estudio aparecen con más frecuencia argumentos en las que se ve a la distribución como una función matemática, se ven argumentación como: “Distribución es la forma o la formula en la que se indica como se acomodan ciertos datos o bases de datos en un espacio ”, en esta parte se percibe aunque con menos fuerza con la que esperábamos la presencia de un conflicto con la concepción antecedente de repartir, dado que en las distribuciones continuas no se

reparte, al menos no de manera puntual, las probabilidades como en el caso continuo. A continuación algunos ejemplos

Nombre (Opcional) : _____

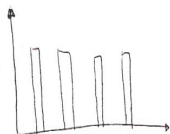
Describe brevemente lo que significa para ti el concepto de distribución
 Lo forma de que una serie de datos son acomodados en alguna situación, por ejemplo teniendo una probabilidad total, que tan probable es que sea alguna situación del total

Como sería para ti una gráfica de una distribución (hacer un dibujo)



Describe brevemente lo que significa para ti el concepto de distribución
 Repartir o dividir en partes iguales.

Como sería para ti una gráfica de una distribución (hacer un dibujo)



Se destaca de estas dos subetapas, por un lado el enriquecimiento de las argumentaciones escritas y gráficas, por ejemplo el número de trazos continuos creció sustancialmente y por otro la permanencia, después de las distribuciones continuas, del argumento “repartir” como concepción de distribución, cuestión que se resume en la tabla 1.

Etapa 0		Etapa 1 - Subetapa 1		Etapa 1 - Subetapa 2	
Idea	Porcentaje	Idea	Porcentaje	Idea	Porcentaje
Agrupar	1.5%	Entregar	1%	Distribuir	1.1%
Enviar	1.5%	Función matemática	1.1%	Frecuencias	1.1%
Llevar	1.5%	Variación	1%	Dividir	2.3%
Organizar	1.5%	Acomodar	3.2%	Función matemática	5.7%
Acomodar	4.6%	Dividir	4.3%	Ordenar / Acomodar	8%
Dar	4.6%	Frecuencias	4.3%	Separar	9.2%
Dividir	4.6%	Distribuir	5.3%	Esparcir / Dispersar	15%
Proporcionar	4.6%	Ordenar	6.4%	Otras	16.1%
Separar	4.6%	Esparcir / dispersar	10.6%	Repartir	19.5%
Ordenar	6.2%	Posibilidad / Probabilidad	11.7%	Posibilidad / Probabilidad	22%
Otras	6.2%	Otras	12.8%	Total	100%
Distribuir	7.7%	Repartir	38.3%		
Repartir	50.8%	Total	100.0%		
Total	100%				

Tabla 1

Subetapa 3. Aquí se organizaron las respuestas para los dos tipos de evaluación, con las que pretendíamos identificar formas más elaboradas de la estructura conceptual, aun reconociendo que hay estudios como el Wilensky (1997); citado por Bakker(2004) que concluyen que distribución es un concepto muy avanzado que, en su completa complejidad, está lejos del alcance del estudiante medio, aunque Bakker sostiene que es posible direccionar el problema desde una situación de nivel informal avanzando mediante un énfasis en la forma, como patrón de variación.

Para las respuestas a la pregunta ¿Qué es la muestra? El 62% lo explica como parte de una población y el 38% lo explica como un conjunto de objetos. Para la pregunta de, ¿Qué significa una distribución muestral? Se distinguieron dos grandes agrupaciones; por un lado los que explican que se trata de una distribución de muestras, 34%, y por otra parte los que ven la distribución muestral como la distribución de los datos de la muestra (el 66%).

Las argumentaciones para la evaluación tipo 2 (operativo) se organizaron clasificando, para cada posible respuesta, las justificaciones que dieron para seleccionar su respuesta, principalmente se basaron en el tamaños de las muestras (Tamaño de muestra) y los que aludieron a conceptos como Teorema del Límite Central, cálculo de probabilidades que genéricamente llamaremos Distribución, aunque el argumento del tamaño de muestra es más frecuente que lo que reportamos; por ejemplo para el inciso c) $P(\bar{x}_1 > 60) = P(\bar{x}_2 > 60)$ se catalogó como Distribución aunque los que la escogieron como respuesta correcta argumentan que como para ambos casos los tamaños de muestra son más de 30, entonces la distribución es normal y por lo tanto las probabilidades son iguales.

Tabla 2

Opción de respuesta	Justificación		Total
	Tamaño de las muestras	Distribución	
a)	33%	11%	44%
b)	17%	15%	32%
c)	0%	8%	8%
d)	11%	6%	17%
Total	61%	39%	100%

ETAPA 2

La clasificación de las respuestas de los alumnos de los grupos de las especialidades es muy similar a las anteriores, con el comentario de que los tipos de gráficos son por lo general continuos (de forma acampanada). Para la pregunta sobre la distribución muestral, el 84% la entiende como una como Distribución de Datos y el 16% como Distribución de Muestras, mejoro la calidad de las argumentaciones pero disminuyo el porcentaje, en relación a la subetapa 3.

Las argumentaciones para la evaluación tipo 2 (operativo) el 22% argumento sobre la Distribución y el 78% justifico su respuesta argumentando sobre el tamaño de muestra, al respecto Bakker (2004) realiza una actividad “una muestra creciente” para que los estudiantes piensen en que ocurre en la grafica al crecer la muestra y otra para razonar acerca de la forma de los datos.

Conclusiones

De los resultados podemos ver las argumentaciones que describen o aluden a la distribución como partición, entrega, etc. y que hemos concretizado como Argumentos cotidianos (los del ámbito cultural) esos argumentos son abandonados por los alumnos que influencia por la clase pasan a usar Argumentos de la Clase, cuestión que se observa con los resultados de las subetapas 1 y 2 donde los Argumentos cotidianos pasaron de un 44% a un 35% aunque se prospecta una disminución en lo futuro dado el 27% de la Etapa 2 (los alumnos de los semestres superiores), digamos que habrá una fracción de alumnos que solo abandonan temporalmente los argumentos del ámbito cultural pero regresaran a ellos, la tabla 1 muestra la evolución de la idea “repartir”; una evolución a la baja y una evolución a la alza en cuanto a ideas como “posibilidad/probabilidad”. Pero por otro lado vemos argumentaciones en las que coexisten ambos tipos, pensamos que esa coexistencia es temporal y paulatinamente cambia al nuevo tipo; a las provistas por la clase aunque no sabemos si fuera del ámbito escolar pueda haber un regreso a las argumentaciones anteriores.

Ubicándonos en el contexto del análisis estadístico de datos se reconoce que un rasgo esencial de ello es que principalmente se trata de describir y predecir características de los conjuntos de datos. Al respecto Bakker y Gravemeijer (2004) nos recuerdan que los estudiantes tienden a concebir un conjunto de datos como una colección de valores individual en lugar de un conjunto de datos que tiene ciertas propiedades y, citando a Cobb(1999), Petrosino, Lehrer y Schauble (2003), enfatizan que los alumnos deben desarrollar la noción de distribución ya que ésta es una estructura conceptual organizativa con la que ellos pueden concebir el conjunto de datos en vez de sólo valores individuales. Al igual que los autores citados podemos concluir que los estudiantes tienden a ver a las muestras como un conjunto de datos individuales y no como un conjunto con ciertas propiedades que vistas individualmente no obtendrán.

En Pfannkuch, M. y Reading, C. (2006) se reta a los investigadores a determinar cuándo se desarrollan las primeras nociones de distribución en los estudiantes y cómo se extienden esas nociones primigenias hacia formas más complejas, así como la necesidad de determinar cómo se conecta el entendimiento de distribución con el de otros conceptos estadísticos y el razonamiento relacionado. Sin duda es un reto que estimula a la búsqueda de respuestas que, también sin duda, impactaran en nuevas estrategias didácticas tendientes a mejorar el aprendizaje de este importante concepto.

Reconocimientos

Los autores agradecen al ITESM el apoyo recibido en la realización de este trabajo a través de la Cátedra de Investigación en Matemática Educativa, Campus Monterrey.

Bibliografía

Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal*, 3(2), 64-83.

Bakker, A. y Gravemeijer, K.P.E (2004). Learning to Reason about Distribution. In D. Ben-Zvi & J. Grafield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking*, Chap. 7, 147-168. Dordrecht: Kluwer Academic Publishers.

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data

analysis. *Mathematical Thinking and Learning*, 1(1), 5–43.

Devore, J. (2004). *Probabilidad y estadística para ingeniería y ciencias* (5ª ed.). México: Thompson.

Montgomery, D. & Runger, G. (1996). *Probabilidad y estadística aplicadas a la ingeniería* (3ª ed.). México: Mc Graw Hill.

Petrosino, A. J., Lehrer, R., & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. *Mathematical Thinking and Learning*, 5(2&3), 131-156.

Pfannkuch, M. & Reading, C. (2006) Reasoning about distribution: A Complex Process. *Statistics Education Research Journal*, 5(2), 4-9.

Watson, Jane M. (2009). The influence of variation and expectation on the developing awareness of distribution. *Statistics Education Research Journal*, 8(1),32-61

Wild, C.J. (2006). The Concept of distribution. *Statistics Education Research Journal*, 5(2),10-26.

Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. *Educational Studies in Mathematics*, 33, 171-202.