

Rincón docente¹

¿Probabilidad para Ingenieros?

Javier Olaya Ochoa

Profesor de la Escuela de Ingeniería Industrial y Estadística-Facultad de Ingeniería -Universidad del Valle.

Introducción

En el fondo la "probabilidad para ingenieros" podría pensarse como una forma de mercadeo de las grandes empresas editoriales. La población de estudiantes de ingeniería es inmensa y la consigna de esta estrategia editorial se entendería como que nada podría ser mejor que llegarles al punto a los estudiantes de ingeniería, sin muchos rodeos. Desde luego, es fácil afirmar que la probabilidad es una sola y que si se requieren libros de probabilidad para ingenieros también deberían escribirse textos de probabilidad para psicólogos, enfermeras, médicos o arquitectos. Según esta idea, el curso de Probabilidad es uno solo. También es fácil contestar a este comentario diciendo que si bien es cierto, también es cierto que los estudiantes de cualquier disciplina tiene un tiempo destinado a cada cosa y que a lo mejor el tiempo que los currículos han reservado a la Probabilidad es lo suficientemente escaso como para que se requieran "adecuaciones" en la docencia de la Probabilidad, posiblemente para tratar de comunicar a los estudiantes lo mejor posible en el marco estrecho de los currículos actuales, enfocados esencialmente a la formación técnica. Esto, que podría llamarse la simplificación de la época, formaría una mezcla "eficiente". Algunos autores de mediados del siglo XX (Parzen y Feller, por

ejemplo) presentaban en sus textos un enfoque más teórico que los textos actuales con algunas incursiones a temas históricos y a aplicaciones que son universalmente válidas. Por ejemplo los temas de genética, los temas de los seguros, los primeros modelos de urnas y en general los procesos estocásticos. El panorama actual reduccionista, impulsado por las visiones curriculares es muy distinto y parece que la "probabilidad para ingenieros" es una propuesta que privilegian el excesivo tecnicismo y la generación de profesionales dirigidos a resolver problemas particulares "del medio". Nuestros ingenieros no parecen estar destinados a generar conocimiento sino a reproducir modelos de adaptación tecnológica ciertamente imaginativos pero en un nivel muy primario, poco creador. En este marco la matemática, la filosofía, la estadística y la probabilidad, para no ir muy lejos, no tienen un espacio muy grande que digamos. Así, la probabilidad como una parte integrante de la cultura de un individuo, no ocupa un lugar en el currículo. Y se genera de esta manera un vacío de formación que en efecto termina siendo reduccionista. El autor explora algunos libros de texto impresos como textos de Probabilidad y Estadística para Ingenieros, tratando de identificar qué pretenden estos autores al incluir en sus textos uno o varios capítulos sobre la Teoría de las Probabilidades. El escrito se refiere al qué incluyen y al cómo lo dicen y aventura algunas opiniones sobre sus impactos sobre la docencia, aclarando que este tema, el de la docencia de la probabilidad para ingenieros, sería objeto de un trabajo diferente.

Notación

Para efectos de este *Rincón Docente* se hará referencia a varios autores. Pero para evitar un exceso de repetición, estos autores se citarán formalmente, con la mayor cantidad de información posible, solo la primera vez que se mencionen. Por ejemplo se citará simplemente a Miller cuando se desee hacer referencia a Miller y Johnson (1999). De hecho esta opción ya se utilizó

¹"Que me dejen el rincón..."

Pete Conde Rodríguez

en la Introducción al referirse a los textos de Parzen (1979) y Feller (1986).

Estadística para Ingenieros

Numerosos textos se han propuesto a lo largo de los años bajo el título aproximado *Probabilidad y Estadística para Ingenieros*. Algunos autores, que podrían llamarse conocidos, son Miller, Walpole (Walpole et al., 1997), Scheaffer (Scheaffer y McClave, 1994), Devore (Devore, 1998) y Ross (Ross, 2002), quienes tienen en común, además de trabajar en Estadística, entre otras cosas, que todos ellos son profesores de universidades americanas. Estos textos tienen la particularidad de incluir uno o varios capítulos iniciales que se refieren a la Teoría de la Probabilidades y de usar al menos el 80% del texto en temas estadísticos. Esto podría entenderse como que estos autores coinciden en que algunas ideas básicas de esta Teoría de las Probabilidades serían convenientes como una fase preliminar al estudio de la Estadística en tanto ciencia aplicable en la Ingeniería. Y que un propósito en sí de la inclusión de ideas probabilísticas en estos textos no sería la enseñanza de la Teoría de las Probabilidades como tal, sino más bien la de sus usos potenciales en la Estadística y de esta en la Ingeniería. Ciertamente o falso este planteamiento, lo que sí es cierto es que así se hace en todos los textos que aquí se refieren. Claro está que en este *Rincón* no se discute la validez de la afirmación anterior, sino que más bien se intenta explorar el cómo y el qué: el qué presentan y el cómo lo presentan, jugando a lo mejor a buscar patrones al estilo Nash (¿Vio la película *Una Mente Brillante*?).

Un modelo discreto de asignación de probabilidades

Todos los autores, ya sea que utilicen o no el título de esta sección de este ensayo, inician su construcción de las ideas fundamentales a través del estudio de un experimento discreto. Miller se apoya en el espacio muestral.

$$S_1 = \{(0,0), (1,0), (0,1), (2,0), (1,1), (0,2)\},$$

en el cual la primera coordenada es el

número de plantas de generación de energía eléctrica, de dos que se construirán, que estarían ubicadas en una región geográfica determinada; y la segunda coordenada indica lo mismo, pero en otra región geográfica. Walpole no va muy lejos y utiliza una moneda y un dado. Scheaffer igual. Ross usa $S_2 = \{f, m\}$ para indicar el género de un recién nacido y $S_3 = \{1, 2, 3, 4, 5, 6, 7\}$ para indicar los números que identifican a siete competidores en una carrera hípica. Devore utiliza una variación del experimento de lanzar tres monedas consistente en probar tres fusibles y registrar si cada fusible probado es defectuoso (D) o no (N), generando el espacio muestral.

$$S_4 = \{DDD, DDN, DND, DNN, NDD, NDN, NND, NNN\}.$$

Otros autores, cuyos textos no necesariamente están dirigidos a ingenieros, siguen igual línea: Mendenhall (Wackerly, Mendenhall y Scheaffer, 2001) usa el mismo dado de Scheaffer [de hecho Scheaffer es coautor del texto de Mendenhall]; Newbold (Newbold, 1998), quien escribe para los Negocios y la Economía, toma prestado el dado de Scheaffer. McClave (McClave, 2000) usa el mismo dado de Scheaffer, con la moneda de Walpole [aunque conviene aclarar que McClave es coautor del texto de Scheaffer]. Triola, en su *Estadística Elemental* (1998) también acude al dado... El texto de Moore (1998) es claramente diferente de los demás en tanto presenta ideas sobre observación, procesamiento y producción de datos antes de requerir de ideas probabilísticas para poder hablar de producir resultados por inferencia. Pero una vez requiere de presentar las ideas básicas de probabilidad... no intenta desprenderse de la moneda, es decir, de un modelo discreto para la asignación de probabilidades.

¿Falta de imaginación? No parece. Más que esto, es una demostración de la facilidad de elaboración que procede al ejemplo introductorio, por cuanto definiciones tales como "experimento aleatorio", "espacio muestral", "punto muestral", "evento simple" y "evento compuesto", se logran a partir de

estos ejemplos simples de manera muy intuitiva. Y las ideas claves en el proceso de asignación de probabilidades se siguen a partir de estas definiciones, de manera expedita: la misma definición de probabilidad introducida por Kolmogorov (Kolmogorov, 2000), las definiciones de probabilidad condicional e independencia, y los teoremas del producto, de la suma, de probabilidades totales y de Bayes. Y en este instante, luego de iniciar en un ejemplo sencillo, es posible desarrollar la idea de variable aleatoria... pero no sin antes enfrentar una etapa intermedia... la asignación de probabilidades en espacios muestrales equiprobables.

Sería interesante en este punto cuestionarse sobre si se trata de una "sistematización" de la enseñanza de las ideas fundamentales en la Teoría de las Probabilidades y si se trata además de promover la enseñanza por repetición excesiva de ejemplos similares al estilo Baldor (Baldor, 1998). Pero sobre este tema ya el profesor Daniel Arbelaez (Arbelaez, 2002) ha discurrecido bastante.

Contar, contar y contar...

Otro tema común en los capítulos sobre la Teoría de las Probabilidades en estos textos pensados (o al menos intitulados) para ingenieros, es el de los métodos de conteo de puntos muestrales. Scheaffer y Mendenhall, quienes parecieran ir de la mano, llaman a estas secciones de sus libros "reglas de conteo útiles en probabilidad" y "herramientas para el conteo de puntos muestrales", respectivamente. Y dentro de su sección siguen el mismo derrotero, que plantea una regla básica ("regla $m \times n$ ") a partir de la cual pueden derivarse todos los métodos de conteo. Miller prefiere llamar a su sección "técnicas de conteo" y Walpole, "conteo de puntos muestrales". Y McClave -quien no escribe para ingenieros- prefiere no mencionar el tema. Ross es un poco más preciso, pues incluye el tema como una subsección ("principios básicos de conteo") de su sección "espacio muestral con resultados igualmente probables". Claro que todos los autores coinciden en la utilidad del conteo en tanto se esté trabajando con espacios muestrales equiprobables y así lo dejan claro

en algún momento. Pero solo Ross, de entre los consultados, lo hace explícito desde la misma organización del texto. Tal vez por esto es que cuando en los cursos de probabilidad para ingenieros se utiliza el experimento de lanzar una moneda dos veces se induce la construcción del espacio muestral.

$$S_5 = \{CC, CS, SC, SS\},$$

en donde un punto muestral consiste en los resultados ordenados del lanzamiento de la moneda. Se tiende entonces a no considerar el espacio muestral, más general,

$$S_6 = \{\{\text{salen dos caras}\}, \{\text{sale una cara y un sello}\}, \{\text{salen dos sellos}\}\}.$$

Una razón para esta aparente preferencia de los docentes podría estar en el hecho que en S_5 los puntos muestrales tienen igual probabilidad de ocurrencia...

Una observación, que podría ser interesante para algunos lectores, es que ninguno de los autores consultados introduce en sus capítulos sobre probabilidad la noción de muestreo con sustitución y, así, no definen esta idea tan cara a los afectos de los profesores de probabilidad. Únicamente Mendenhall, en su despedida del capítulo sobre definiciones básicas, menciona la importancia para el Muestreo de definir claramente cómo se selecciona la muestra, en atención al hecho cierto que tal situación afectará la asignación de probabilidades. Así que, concluye Mendenhall, el procedimiento de Muestreo se debe describir claramente si se desea construir inferencias válidas sobre una población a partir de una muestra. Siguiendo a Ross, la definición de sustitución podría aparecer en alguna parte de la subsección sobre métodos de conteo, en la sección de espacios equiprobables... pero ni el mismo Ross, ni ninguno de los demás autores consultados, lo hace. Devore, por ejemplo, presenta definiciones fundamentales como la misma definición de probabilidad (desde luego sin apartarse de Kolmogorov), sin mencionar los espacios equiprobables ni, como es de suponer, los métodos de conteo. Finalmen-

te dedica unas breves líneas a estas estrategias de conteo, sin mencionar el muestreo con reemplazo.

El autor de este *Rincón* ha afirmado públicamente que una definición de muestreo con sustitución no es necesaria para la enseñanza de las ideas básicas en probabilidad para ingenieros, arrastrado posiblemente por la forma como estos autores abordan sus capítulos sobre la Teoría de la Probabilidades. Esto podría no ser cierto, en particular por la importancia mencionada por Mendenhall. Sin embargo, tal afirmación podría aparecer razonable si se mira la definición del muestreo con sustitución como un problema particular del conteo en el marco de los espacios muestrales equiprobables, los cuales pueden no ser los más comunes, aunque sí los más útiles para presentar las ideas básicas, tal como se mencionó. Y para presentar las ideas básicas, estos autores no parecen precisar de esta idea tan importante. El autor de este *Rincón*, de hecho, jamás la ha utilizado en sus cursos básicos de Estadística, posiblemente por considerar, siguiendo a Ross, que son extensiones elementales del conteo en espacios muestrales equiprobables.

Funciones de valor real...

El siguiente paso (¿natural?) es moverse de las definiciones fundamentales en el proceso de asignación de probabilidades al concepto más avanzado de variable aleatoria. Aquí la construcción del modelo discreto para asignación de probabilidades juega un papel crucial, por cuanto una variable aleatoria discreta surge de manera sencilla de un espacio muestral finito, a partir de la definición de un "evento numérico". Por ejemplo, si ya no es importante observar y registrar el resultado de lanzar una moneda dos veces, si no observar y registrar el número de caras que se obtienen al lanzar una moneda dos veces, entonces estamos ante un experimento cuyos resultados son "eventos numéricos". En este caso el espacio muestral.

$$S_6 = \{\{\text{salen dos sellos}\}, \{\text{sale una cara y un sello}\}, \{\text{salen dos caras}\}\},$$

se convierte en

$$S_7 = \{0, 1, 2\}.$$

La mayoría de los autores consultados utilizan la misma estrategia para introducir el concepto de variable aleatoria y definen de manera más o menos formal una variable aleatoria diciendo, palabras más, palabras menos, "... es una función que proyecta eventos de un espacio muestral sobre la recta real" (Ang y Tang, 1975). Esta definición ratificaría la importancia de definir un espacio muestral, por lo que resulta sencillo pasar de la definición a su comprensión. Si X es el mecanismo de asignar un evento numérico de S_7 a un evento de S_6 , de tal manera que cada evento simple de S_6 sea enviado a uno y solo uno de los eventos numéricos de S_7 , entonces X es una variable aleatoria. Newbold se separa ligeramente y arriesga su propia definición al afirmar que una variable aleatoria "... es una variable que toma valores numéricos determinados por el resultado de un experimento". Ya el lector se formará sus propias ideas sobre esta definición. Ross prefiere una presentación poco formal del tema y evita la introducción de una definición. Esta es una aproximación muy interesante, más aún viniendo de Ross, un autor muy conocido por sus excelentes aportes en Probabilidad, en Procesos Estocásticos y en Estadística Matemática. Cuál sea la mejor vía parece ser una pregunta abierta...

Ahora bien, no todas las variables son discretas. Así que aún queda por resolver una especie de lío... introducir la idea de variable aleatoria continua. Aquí hay una diversidad de estrategias para formalizar este concepto y hacerlo llegar al ingeniero. Miller decide conducir una larga disquisición sobre la "idea" de variable aleatoria continua que finaliza en una afirmación formal: que si la derivada $f(x)$ de la función de distribución $F(x)$ existe y es continua, entonces $f(x)$ es una densidad. Esto implica que las variables aleatorias discretas ya se han utilizado para definir una función de distribución, algo que no todos los autores hacen. Ross introduce el tema diciendo que si una variable aleatoria toma un número finito o contable de valores posibles,

se le llama discreta. Y que la que toma un continuo de valores se le llama continua. Pero a la larga regresa a una definición similar a la de Miller que conduce, como puede esperarse, a igual conclusión. Mendenhall usa una opción ligeramente diferente, al iniciar desde la función de distribución y pasar a una definición de variable aleatoria continua diciendo que si la función de distribución es continua, entonces la variable aleatoria es continua. Aclara luego que la derivada de la función de distribución debe existir y que debe ser continua excepto quizás en un número finito de puntos en cualquier intervalo finito. Curiosamente, Scheaffer, en su texto para ingenieros, no se aparta mucho de Miller ni de Ross, más aún cuando la idea de su texto conjunto con Mendenhall parece ser muy intuitiva y podría "llegarle" fácilmente a los ingenieros.

Walpole asocia variables aleatorias discretas con espacios muestrales discretos, como se mencionó. Y asocia variables aleatorias continuas con espacios muestrales continuos. El problema es que estos espacios muestrales no son tan sencillos de "ver" como los primeros, y nociones simples en un espacio muestral discreto pueden llegar a ser complejas en uno continuo. Por ejemplo la definición de un "punto muestral" en un espacio muestral continuo es ya de por sí suficientemente compleja. Luego define una función de densidad, seguida de la definición de función de distribución. Es decir, exactamente lo contrario que Mendenhall.

Ross, Miller y Walpole definen estos dos tipos de variables aleatorias en la misma sección de sus textos. Scheaffer y Mendenhall lo hacen en secciones separadas... y a partir de este punto todos siguen en esencia la misma ruta...

A modo de conclusión

Siguiendo la definición de variable aleatoria discreta está la de distribución de probabilidad y en seguida el de valores esperados, aunque resulta interesante anotar que Miller presenta la noción de esperanza matemática sin mencionar las distribuciones de probabilidad. Las mismas definiciones para

el caso continuo siguen a la definición de función de densidad. Todos los autores finalizan sus capítulos sobre la Teoría de las Probabilidades mencionando algunas distribuciones y densidades.

Los textos consultados, aunque actuales, son pocos como para atreverse a formular ideas de tipo general. Desde una mirada bien Estadística, se trataría de construir inferencias a partir de una muestra inadecuada. Por ejemplo, es posible que otros autores no consultados traten estos temas de Probabilidad para Ingenieros de manera diferente. En particular habría sido interesante incluir el texto del profesor Vardeman (Vardeman, 1993), con su aproximación a la Estadística para Ingenieros desde la óptica del Diseño Experimental, que muy posiblemente ha asumido Nelson (Nelson, Coffin y Copeland, 2003). También habría sido interesante consultar autores no americanos, pero no tuve acceso a un texto de Estadística pensado para ingenieros que no fuese de autores de esta nacionalidad o que trabajen en los Estados Unidos de América. Y vale la pena mencionar, para terminar, que este *Rincón* se refiere a los contenidos de los textos y al cómo los presentan sus autores. Pero no al cómo se debería practicar la docencia de la Probabilidad para Ingenieros. Este ciertamente, es un tema bien distinto, que espero motive a alguno de los eventuales visitantes de este *Rincón*... En particular, el autor preferiría disponer de más tiempo para dedicarle a la enseñanza de la probabilidad a los estudiantes de ingeniería, pero reconoce como muy baja la probabilidad de que las tendencias curriculares actuales le permitan más tiempo para ello... El qué se debe hacer y el cómo se debe hacer son entonces preguntas abiertas, cuyas respuestas deberán considerar las restricciones reales de tiempo disponible para esta actividad específica en los currículos de ingeniería. Lo que no es claro es si los autores de estos textos han considerado este hecho en su formulación, o si simplemente quieren "ir al punto"... Y en este punto cada lector deberá formarse su propia impresión.

Agradecimientos: El autor desea agradecer los comentarios y sugerencias de los profesores Daniel Arbeláez y Rafael Klinger.

Bibliografía

- ANG, A. H. S. y Tang, W. H. (1975) *Probability Concepts in Engineering Planning and Design, Volume 1 - Basic Principles*, Wiley.
- ARBELAEZ, D. (2000) *Mi Amargo IQ. El Malpensante*, 27:46-51.
- BALDOR, A (1998) *Algebra*, Publicaciones Cultural
- DEGROOT, M. y Schervish, M. J. (2001) *Probability and Statistics*, 3Ed, Pearson.
- DEVORE, J. L. (1998) *Probabilidad y Estadística para Ingeniería y Ciencias*, 4Ed, Thomson.
- FELLER, W. (1986) *Introducción a la Teoría de la Probabilidades y sus Aplicaciones Volumen 1*, Limusa.
- KOLMOGOROV (2000) A. N. *Foundations of the Theory of Probability*, American Mathematical Society.
- MILLER, I. y Johnson, R. A. (1999) *Miller and Freund's Probability and Statistics for Engineers*, 6Ed, Pearson.
- MONTGOMERY, D. C. y Runger, G. C. (2002) *Applied Statistics and Probability for Engineers*, 3Ed, Wiley.
- MOORE, D. S. y McCabe, G. P. (1998) *Introduction to the Practice of Statistics*, 3Ed, Freeman.
- NELSON, P., Coffin, M. y Copeland K. A. F. (2003) *Introductory Statistics for Engineering Experimentation*, Elsevier.
- PARZEN, E. (1979) *Teoría Moderna de las Probabilidades y sus Aplicaciones*, Limusa
- ROSS, S. M. (2002) *Probabilidad y Estadística para Ingenieros*, 2Ed, McGraw-Hill.
- SCHEAFFER, R. L. y McClave, J. T. (1994) *Probability and Statistics for Engineers*, 4Ed, Wadsworth.
- TRIOLA, M. F. (1998) *Elementary Statistics*, 7Ed, Addison Wesley.
- WACKERLY, D. D., Mendenhal, W. y Scheaffer, R. L. (2001) *Mathematical Statistics with Applications*, 6Ed, Brooks/Cole.
- WALPOLE, R. E., Myers, R. H., Myers, S. L. y Myers, S (1997) *Probability and Statistics for Engineers and Scientists*, 6Ed, Prentice Hall.
- VARDEMAN, S. (1993) *Statistics for Engineering Problem Solving*, Brooks/Cole.