



La importancia de los supuestos en la creación y resolución de problemas en la teoría de probabilidad.

Manuel González Navarrete
manuelg@ime.usp.br
Instituto de Matemática e Estatística
Universidade de Sao Paulo, Brasil.

Recibido: Abril, 17 2016 Aceptado: Julio 20, 2016

Resumen. En este trabajo proponemos un enfoque para abordar los conceptos básicos de la teoría de probabilidad en la formación de profesores de matemática. Basados en la resolución de problemas, presentamos y ejemplificamos las nociones que fundamentan la teoría moderna de probabilidad; σ -álgebra, variable aleatoria y distribución de probabilidad. Nuestra propuesta busca realzar la importancia de los supuestos en el planteamiento y resolución de problemas del área de probabilidades. Como referencia usamos una situación propuesta por U. Malaspina [19], en este estudio analizamos un par de problemas matemáticos que derivan de dicha situación.

Palabras clave: Resolución de problemas, axiomas de probabilidad, variable aleatoria, distribución de probabilidad, didáctica de la probabilidad

Abstract. We propose an approach to address the concepts of probability theory in mathematics teachers training. Based on problem-solving, we present the concepts underlying the modern probability theory; σ -algebra, random variable and probability distribution. Our proposal seeks to highlight the relevance of assumptions in the creation and solution of problems in the area of probability. We use as reference a situation proposed by U. Malaspina [19]. This work analyzes a couple of problems that derive from that.

KeyWords: Problem-solving, probability axioms, random variable, probability distribution, teaching of probability

1.1 Introducción

La teoría de probabilidad, como actualmente la conocemos, debe sus orígenes a diversos matemáticos del siglo XVII, tales como Pierre Fermat, Blaise Pascal, Christiaan Huygens, Jacob Bernoulli y Abraham de Moivre. Aunque, específicamente la llamada teoría clásica de probabilidad se estructura en los trabajos de Pierre Simon, Marqués de Laplace (1749-1827). En las obras *Théorie Analytique des Probabilités* [15] y *Essai philosophique sur les probabilités* [16], Laplace presenta una exposición sistemática del cálculo de probabilidades, que hasta entonces solo empleaba los recursos de la aritmética. De esta forma, Laplace es capaz de organizar los desarrollos obtenidos durante un siglo y medio, exponiendo una formulación, relativamente definitiva, del cálculo de probabilidades, tal como evidencia el trabajo de Vargas [36].

Sin embargo, la rigurosidad matemática de la actual teoría de probabilidad se alcanzó un siglo más tarde, en la obra del matemático soviético Andréi Kolmogorov (1903-1987); escrita en alemán y titulada *Grundbegriffe der Wahrscheinlichkeitsrechnung*, traducida posteriormente como *Foundations of the Theory of Probability* [14]. El trabajo de Kolmogorov propone los axiomas que sustentan la llamada teoría moderna de probabilidad. Esta axiomática es basada en la teoría de conjuntos y la teoría de la medida, desarrollada años antes por Lebesgue, Borel y Frechet entre otros. Citamos como referencia didáctica los trabajos de Jiménez Saavedra [11], Debnath y Basu [5] y Shafer y Vovk [34].

Las aplicaciones de esta teoría moderna extrapolan a todas las áreas del conocimiento humano. A modo de ejemplo, mencionamos los llamados procesos estocásticos, con utilidad en la caracterización de los procesos de interacción molecular, el análisis matemático del mercado financiero, el estudio de la evolución de especies y la construcción de modelos para predicciones climáticas (ver Rincón [29] y Ross [31]). Es claro que la sociedad actual se ha visto influenciada por el desarrollo de métodos rigurosos para realizar predicciones probabilísticas. De esta forma, es importante para la mayoría de los ciudadanos comprender cómo se hacen las asignaciones de probabilidad, y cómo estas contribuyen en la toma de decisiones.

En este sentido, varios autores resaltan la importancia de que los individuos se familiaricen con estos conceptos para comprender y predecir mucho mejor el mundo en que vivimos (ver, por ejemplo, Jimenez y Jimenez [10] y Jones [12]). De otra forma, se dice que las aplicaciones de esta teoría hacen que su enseñanza sea un deber de los sistemas educativos actuales. A pesar de que por mucho tiempo esta área de la matemática no figurara en los planes de estudio de las carreras universitarias, mucho menos en los de la enseñanza elemental y media (Santaló [32]).

Motiva el hecho de que en la última década diversos estudios han sido dedicados a la didáctica de la probabilidad, entre estos trabajos: Batanero et al. [3], González Navarrete [8], Klymchuk y Kachapova [13], Ollerton [23] y Viali y Cury [37]. Sin embargo, es importante mencionar el análisis realizado por Alsina y Vásquez [2] sobre los conocimientos probabilísticos en un grupo de profesores de educación primaria en Chile. La conclusión principal es clara, resulta necesaria una mayor especialización del profesorado en todas las facetas de su conocimiento didáctico-matemático del contenido de probabilidades.

De forma complementaria, se busca seguir los enfoques constructivistas que resaltan la importancia de la resolución de problemas en la didáctica de la matemática (entre ellos, Lesh y Doerr [17], Nunokawa [22], Sepúlveda et al. [33] y Taplin y Chan [35]). En este sentido, O'Shea y Leavy [24] plantean que la

resolución de problemas matemáticos es parte esencial de la obra de cualquier matemático y para llegar a ser un matemático uno debe convertirse en un solucionador de problemas. Para que esto suceda, los estudiantes deben participar en la solución de problemas reales en las aulas para convertirse en solucionadores de problemas también en la vida cotidiana.

Por dichas razones, este artículo presenta un enfoque teórico para abordar los contenidos de la teoría de probabilidad en los cursos de formación de profesores secundarios de matemática, destacando la importancia de los supuestos en la creación y resolución de problemas probabilísticos.

Particularmente, interesa la resolución de un problema de probabilidad enunciado por el Prof. Uldarico Malaspina [19]. Basados en dicho problema serán abordados los conceptos de la teoría moderna de probabilidad, a través de ejemplos. De esta forma, se enuncian los conceptos de σ -álgebra, variable aleatoria, distribución de probabilidad e independencia de variables aleatorias. Será visto como los supuestos influyen en la resolución de un problema de probabilidad. Además, se busca discutir el desafío que supone la creación de problemas matemáticos por parte de profesores secundarios y de los estudiantes de pedagogía, tanto en su formación como en su futuro quehacer profesional (ver Malaspina [20]).

El principal aporte de esta propuesta es realzar la importancia de los supuestos en el modelamiento de la aleatoriedad. De otra forma, mostraremos mediante ejemplos, que los supuestos asumidos como punto de referencia para proponer un modelo probabilístico, pueden tornar el problema en una trivialidad. Así como también, veremos que suposiciones más realistas enriquecen la resolución del problema, haciendo más complejas las herramientas matemáticas necesarias para enfrentarlo. En esta línea, siguiendo a Francisco y Maher [7], proporcionar a los estudiantes la oportunidad de trabajar en tareas complejas en comparación con tareas simples es crucial para estimular su razonamiento matemático.

El artículo se estructura de la siguiente manera: en la Sección 1.2 presentamos la motivación de nuestro trabajo. La Sección 1.3 incluye los conceptos de σ -álgebra, variable aleatoria y distribución de probabilidad. En la Sección 1.5, abordamos la resolución del problema bajo los supuestos de independencia y de distribución idéntica. Como veremos, en Malaspina [19] estos supuestos fueron implícitamente asumidos, lo que hace que el problema sea trivial. En la Sección 1.6, para incluir interés matemático a la situación en estudio, proponemos dos nuevas interrogantes eliminando la suposición de distribuciones idénticas. De esta forma, discutimos algunos modelos de probabilidad usualmente mencionados en la literatura. En la Sección 1.8 resolvemos nuestras preguntas bajo ciertos supuestos distribucionales. La Sección 1.9 contiene las conclusiones de nuestro trabajo y algunos comentarios para profundizar los contenidos del área de probabilidad.

1.2 Motivación

La motivación para este trabajo proviene de un problema de probabilidades discutido por U. Malaspina [19] en la sección *El rincón de los problemas* de la Revista Iberoamericana de Educación Matemática UNIÓN.

En dicho artículo se relata una experiencia de creación de problemas realizada con estudiantes de pedagogía en matemática. Los estudiantes, separados en grupos, se enfocaron en la creación de problemas matemáticos basados en la siguiente situación:

“En un sistema vigesimal de calificación, las notas de un alumno en los dos primeros exámenes de los cursos A y B son 8 y 12, y 10 y 14 respectivamente. Las notas finales se obtienen como media aritmética de tres exámenes, los cursos se aprueban con notas mayores o iguales a 11 y procede el redondeo habitual al segundo decimal.”

Como es de esperar, la mayoría de los problemas propuestos por los estudiantes buscaba la nota mínima necesaria para obtener determinado promedio final en cada curso. Esto es, problemas matemáticos que pueden ser resueltos mediante ecuaciones (o inecuaciones) lineales. En este sentido, si respondemos a la pregunta *¿qué notas puede obtener el alumno en su tercer examen para aprobar el curso A?*, denotando x el valor de la tercera nota, debemos resolver,

$$\frac{8 + 12 + x}{3} \geq 10.95. \quad (1.1)$$

Esto es, para aprobar el curso A el valor de x debe satisfacer,

$$x \geq 12.85. \quad (1.2)$$

De manera análoga, para aprobar el curso B, cuando denotamos por y el valor de la tercera nota, se necesita

$$y \geq 8.85. \quad (1.3)$$

Sin embargo, hubo un grupo que planteó el problema de probabilidades que nos motiva. A pesar de que dichos estudiantes no presentaron una solución matemática, para Malaspina la discusión generada resultó valiosa pues el tema de probabilidades es muy poco tratado en los estudios secundarios y aún en los cursos básicos de la educación superior.

La pregunta analizada en Malaspina [19] es la siguiente,

¿En qué curso hay mayor probabilidad de que el alumno obtenga nota final aprobatoria?

En [19] se incluye la discusión de algunas soluciones basadas en la probabilidad clásica de Laplace. Destacamos el esfuerzo de Malaspina de profundizar los conceptos de variable aleatoria y distribuciones de probabilidad, pero debemos indicar que existen imprecisiones a la hora de asumir los supuestos y resolver la interrogante planteada. Por este motivo, parte de nuestro trabajo consiste en aclarar la implicancia de las suposiciones que sean asumidas y ejemplificar la utilidad de los conceptos de probabilidad en el modelamiento de la aleatoriedad.

A seguir, enunciamos algunos comentarios sobre este problema.

- Es importante comprender que el objeto de interés son las notas del tercer examen en el curso A y curso B. De otra forma, denotamos las variables (aleatorias) X e Y representando la nota (desconocida) del 3er examen en los cursos A y B, respectivamente. Así, tabulamos los datos en la Tabla 1.1.

	Curso A	Curso B
1er Examen	8	10
2do Examen	12	14
3er Examen	X	Y

Tabla 1.1: Resumen de notas. Las notas desconocidas X e Y serán definidas como variables aleatorias.

- Para responder a la pregunta debemos asignar distribuciones de probabilidad a las variables X e Y . De la forma en que fue formulada, esta interrogante deja abierta la opción de asumir diferentes supuestos a la hora de resolverla. Por ejemplo, se puede asumir que las notas de ambos cursos son independientes. Otro factor a considerar son las distribuciones de probabilidad que rigen las notas de cada curso, o variables aleatorias X e Y (ver Sección 1.3).
- Al observar que las notas en el curso B han sido más altas en ambas pruebas, es razonable pensar que el curso con mayor probabilidad de ser aprobado sea el curso B. De hecho, como veremos en la Sección 1.5, los supuestos implícitamente asumidos en el trabajo de Malaspina [19] tornan esta una respuesta trivial.
- Para comprender la afirmación anterior basta entender los conceptos básicos de la teoría de probabilidad propuesta por Kolmogorov (Sección 1.3.2). De otra forma, un profesor de matemática en conocimiento de los axiomas de probabilidad sería capaz de formular otro tipo de interrogantes, entendiendo que esta pregunta no es interesante a la hora de ser resuelta matemáticamente.

De esta manera, nuestro objetivo es destacar la importancia de los supuestos asumidos a la hora de resolver, y más importante, crear un problema de probabilidad. Analizamos los supuestos en las resoluciones presentadas por Malaspina [19]. Además, planteamos un par de nuevas interrogantes para ser resueltas en base a otras suposiciones, para así tornar más interesante la matemática involucrada.

Finalmente, afirmamos que comprender la importancia de los supuestos implica un apoderamiento didáctico por parte del docente, que le permitirá tener la capacidad de crear y resolver problemas de probabilidades con la rigurosidad matemática necesaria.

Por lo tanto, de forma general buscamos aportar a la formación de profesores de matemática y particularmente al conocimiento didáctico-teórico del contenido de probabilidades.

1.3 Conceptos básicos de la teoría de probabilidad

En esta sección abordamos los conceptos y definiciones que sustentan la teoría de probabilidad. Esta breve introducción puede ser complementada con los textos de Magalhães [18], Montes [21] y Rincón [28]. El contenido de esta sección será aplicado en las secciones siguientes, en las cuales bajo diferentes supuestos se dará respuesta al problema que nos motiva. Además, en la Sección 1.9 comentamos algunas formas de profundizar los contenidos matemáticos asociados a la probabilidad moderna.

La teoría matemática conocida como probabilidad debe sus inicios a los intentos de Gerolamo Cardano de analizar las chances en juegos de apuesta. Posteriormente, en el siglo XVII los matemáticos Pierre de Fermat y Blaise Pascal intercambian correspondencias en las cuales discuten algunos problemas de probabilidad. En la secuencia, en el año 1657, Christiaan Huygens publicó un libro sobre este asunto. La llamada teoría clásica de probabilidad queda cimentada en los trabajos de Laplace, en el siglo XIX (ver [5] y [36]).

El desarrollo de esta teoría culminó en lo que conocemos como teoría moderna de probabilidad. Su fundamento axiomático fue dado en 1933 por Andréi Kolmogorov, quién combinó la noción de espacio muestral con la teoría de la medida. Rapidamente esta definición se convirtió en la base axiomática de

la actual teoría de probabilidad [11, 34].

En la secuencia, completamos el estudio del problema introduciendo los conceptos de probabilidad que son necesarios para interpretar y resolver la interrogante que nos motiva.

1.3.1 Experimento aleatorio y espacio muestral

Para comenzar debemos definir lo que se conoce como *experimento aleatorio*. En un experimento aleatorio entra en juego el factor incertidumbre o azar.

Un fenómeno o experimento aleatorio es aquel que al ser repetido bajo las mismas condiciones iniciales, puede presentar resultados diferentes. De otra forma, aunque no pueda ser repetido, al realizar una experiencia particular no es posible predecir el resultado exacto. Este concepto es opuesto a la idea de fenómeno *determinístico*, en el que conociendo los factores iniciales del experimento podemos describir con exactitud el resultado del mismo.

Todo experimento aleatorio debe ser asociado a un conjunto llamado *espacio muestral*, que denotamos por Ω , este se refiere al conjunto que agrupa todos los posibles resultados que un experimento aleatorio pueda tener. Como veremos en ejemplos, para un mismo experimento aleatorio es posible definir diferentes espacios muestrales, esto dependiendo de la interrogante que deseamos plantear. Este hecho es importante de ser clarificado, porque el espacio muestral debe ser adecuadamente definido para que las conclusiones probabilísticas sean correctas. Así, una vez establecido un espacio muestral correctamente, la rigurosidad matemática está garantizada por los axiomas de probabilidad (Definición 1.2).

A seguir, presentamos algunos ejemplos de experimentos aleatorios (EA) y sus espacios muestrales (Ω) asociados.

Ejemplo 1.1

Un ejemplo bastante común en los textos de probabilidad.

EA: Lanzar un dado dos veces.

$$\Omega = \{(1,1);(1,2); \dots, (6,5);(6,6)\}.$$

Ejemplo 1.2

Una secretaria necesita realizar una llamada, pero olvida los últimos 2 dígitos del teléfono a marcar. Si decide completar el número telefónico de manera aleatoria.

EA: Discar dos dígitos al azar.

$$\Omega = \{00, 01, \dots, 10, 11, \dots, 98, 99\}$$

Ejemplo 1.3

El juego del "cachipún" (también conocido como piedra, papel o tijera) es un buen ejemplo de experimento aleatorio.

EA: Jugar una partida de cachipún.

$$\Omega = \{ (Piedra, Piedra) ; (Piedra, Papel) ; (Piedra, Tijera) ; (Papel, Piedra) ; (Papel, Papel) ; (Papel, Tijera) ; (Tijera, Piedra) ; (Tijera, Papel) ; (Tijera, Tijera) \}.$$

Ejemplo 1.4

En una fábrica de ampollitas están interesados en estudiar el tiempo de vida de sus productos. Se enciende una ampollita y se espera hasta el instante que deja de funcionar. Asuma que el tiempo es una variable continua (pertenece a los números reales positivos, \mathbb{R}^+).

EA: Analizar el tiempo de vida de una ampollita.

$$\Omega = [0, \infty) \subset \mathbb{R}.$$

Ejemplo 1.5

La nota obtenida en un examen depende de varios factores, por ejemplo: el tiempo y forma de estudio, la dificultad de las preguntas o el estado de ánimo del estudiante. De esta forma, parece razonable asumir que la nota de un examen es aleatoria, pues no podemos predecirla con certeza.

EA: Hacer el tercer examen del curso A.

Observación: Para este EA podemos definir el espacio muestral de varias formas. Es claro que los posibles resultados deben representar todas las notas que pueden ser obtenidas en el examen. Note también que se usa un sistema de notas poco usual, escala de 0 a 20, pero se sigue el enunciado propuesto en [19]

Aquí consideramos dos posibles formas de definir el espacio muestral:

1. Un sistema de notas con un valor decimal,

$$\Omega' = \{0.0; 0.1; \dots; 20\} := \{0 + 0.1 \cdot n \mid n = 0, 1, 2, \dots, 200\}.$$

2. Asumiendo las notas como valores continuos, permitiendo mayor flexibilidad y mejorando las aproximaciones,

$$\Omega = [0, 20] \subset \mathbb{R}. \tag{1.4}$$

1.3.2 Definición axiomática de probabilidad

A continuación, presentamos un concepto de la teoría de conjuntos que da fundamento a la definición axiomática de probabilidad. De esta manera, lo que haremos es caracterizar una clase de subconjuntos de nuestro espacio muestral Ω . Esta clase, llamada σ -álgebra, será usada para la definición de probabilidad dada por Kolmogorov.

Definición 1.1 (σ -álgebra)

Una clase de subconjuntos de Ω , representada por \mathcal{F} , es denominada una σ -álgebra si satisface las siguientes propiedades:

- (a) $\Omega \in \mathcal{F}$,
- (b) Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$,
- (c) Si $A_i \in \mathcal{F}, i \geq 1$, entonces $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

En palabras, el espacio muestral debe pertenecer a la clase y si algún conjunto está incluido, su complemento también debe estarlo. Además, para cualquier colección (finita o) infinita contable de elementos dentro de la clase \mathcal{F} , su unión debe estar incluida en dicha clase.

Ejemplo 1.6

De manera simple suponga que $\Omega = \{1, 2, 3\}$. Es fácil probar que las siguientes clases de subconjuntos son σ -álgebras.

$$\mathcal{F}_1 = \{\emptyset, \Omega\}$$

$$\mathcal{F}_2 = \{\emptyset, \{1\}, \{2,3\}, \Omega\}$$

$$\mathcal{F}_3 = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \Omega\}$$

La clase \mathcal{F}_1 es la menor σ -álgebra que podemos construir. Mientras que la clase \mathcal{F}_3 es conocida como conjunto potencia de Ω , la cual es la mayor σ -álgebra de Ω . Las clases \mathcal{F}_1 y \mathcal{F}_3 son llamadas σ -álgebra trivial y discreta, respectivamente.

El siguiente ejemplo es importante en el desarrollo de la teoría de probabilidad. Específicamente, cuando definimos variables aleatorias continuas, que veremos en la Definición 1.6, usamos la σ -álgebra a seguir.

Ejemplo 1.7 (σ -álgebra de Borel)

Si consideramos $\Omega = \mathbb{R}$, podemos definir una clase de subconjuntos que llamamos la σ -álgebra de Borel.

Esta se define como la mínima σ -álgebra sobre \mathbb{R} que contiene a los subconjuntos cerrados, a los intervalos abiertos (a,b) o cerrados $[a,b]$, a los intervalos semiabiertos de la forma $(a,b]$, o a los intervalos de la forma $(-\infty, b]$.

Formalmente, el siguiente paso es definir medidas que cuantifiquen la incertidumbre sobre los posibles resultados de un experimento aleatorio. En otras palabras, asociaremos valores que representen la

probabilidad de ocurrencia de cada uno de los conjuntos que componen la σ -álgebra correspondiente. En general, los subconjuntos que están en la σ -álgebra \mathcal{F} son denominados *eventos* y únicamente a ellos se atribuye probabilidad.

La teoría de probabilidad fundada en los años 1930 por A. N. Kolmogorov, presenta un conjunto de axiomas que una función, con valores en $[0,1]$, debe satisfacer para ser llamada de *probabilidad*.

Definición 1.2 (Probabilidad)

Sea \mathcal{F} una σ -álgebra de subconjuntos de Ω . Una función $\mathcal{P} : \mathcal{F} \rightarrow [0,1]$, es una probabilidad si satisface los *Axiomas de Kolmogorov*:

(A1) $\mathcal{P}(\Omega) = 1$,

(A2) Para todo subconjunto $A \in \mathcal{F}$, $\mathcal{P}(A) \geq 0$,

(A3) Para toda secuencia $A_1, A_2, \dots \in \mathcal{F}$, mutuamente excluyentes, tenemos

$$\mathcal{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathcal{P}(A_i). \tag{1.5}$$

Recordamos que dos eventos A y B son *mutuamente excluyentes* si, y solamente si, $A \cap B = \emptyset$. Además, considerando (A1)-(A3) es posible probar que $0 \leq \mathcal{P}(A) \leq 1$, para todo evento $A \in \mathcal{F}$. Formalmente, el trío $(\Omega, \mathcal{F}, \mathcal{P})$ es llamado de *espacio de probabilidad*.

Entiéndase que cuando hablamos de probabilidad, matemáticamente estamos hablando de valores entre 0 y 1, los que representan la incerteza (o creencia) sobre posibles resultados de un experimento aleatorio. En algunas ocasiones se habla erróneamente de probabilidad como un valor porcentual, conengamos que eso puede ser llamado de chance o posibilidad. Por ejemplo, se puede decir que mañana hay 70% de chance de lluvia (o la probabilidad de lluvia es 0.7).

Precisamos aclarar la forma en que estos axiomas generalizan las ideas de probabilidad de Laplace, las que conocemos como teoría clásica de probabilidad. Entonces, en el caso contable finito, la probabilidad clásica nos dice que para cada $A \subset \Omega$:

$$P(A) = \frac{\text{Número de elementos en } A}{\text{Número total de elementos en } \Omega}. \tag{1.6}$$

Es fácil comprobar que la probabilidad clásica satisface los axiomas (A1)-(A3). Sin embargo, esta definición se refiere a subconjuntos unitarios equiprobables. Observe que dos conjuntos con el mismo número de elementos tendrán la misma probabilidad (1.6). Este supuesto restringe bastante las posibilidades de aplicación de la probabilidad de Laplace, como veremos en la Sección 1.7.

Por otra parte, note el problema que se produce cuando el conjunto Ω es infinito, y si es aún infinito no contable como $\mathbb{R}^n, n \geq 1$. Con estas situaciones es posible entender como la formulación clásica pierde rigor matemático en problemas más generales. En este sentido, la teoría basada en los Axiomas (A1)-(A3) es responsable por la construcción rigurosa de los conceptos de probabilidad moderna. Los elementos básicos de esta teoría serán expuestos a seguir.

Para finalizar este apartado, probaremos algunas propiedades de la (función) probabilidad dada en la Definición 1.2.

Proposición 1.1

Dado $(\Omega, \mathcal{F}, \mathcal{P})$, considere $A, B \in \mathcal{F}$,

$$(P1) \quad \mathcal{P}(A^c) = 1 - \mathcal{P}(A),$$

(P2) Sean A y B dos eventos cualquiera, entonces

$$\mathcal{P}(B) = \mathcal{P}(B \cap A) + \mathcal{P}(B \cap A^c), \quad (1.7)$$

(P3) Si $A \subset B$, entonces $\mathcal{P}(A) \leq \mathcal{P}(B)$.

Demostración:

(P1) Los eventos A y A^c son mutuamente excluyentes. Además, $A \cup A^c = \Omega$, luego por (A1) y (A3),

$$1 = \mathcal{P}(\Omega) = \mathcal{P}(A \cup A^c) = \mathcal{P}(A) + \mathcal{P}(A^c). \quad (1.8)$$

(P2) Para dos eventos cualquiera A y B , siempre es posible escribir $B = (B \cap A) \cup (B \cap A^c)$. Esta unión es disjunta, entonces por (A3) la propiedad es válida.

(P3) Si $A \subset B$, entonces podemos escribir, $B = A \cup (B \cap A^c)$. Así, como la unión es disjunta usamos (A3) y dado (A2), tenemos

$$\mathcal{P}(B) = \mathcal{P}(A \cup (B \cap A^c)) \geq \mathcal{P}(A). \quad (1.9)$$

En la secuencia, definimos algunos elementos importantes en la teoría de probabilidad, y en particular en nuestro estudio. Esto es, los conceptos de variable aleatoria y función de distribución.

1.3.3 Variables aleatorias, función de distribución y esperanza matemática

En general, cuando se desarrolla un experimento aleatorio, el observador está interesado en uno o más valores (cantidades) asociados a las ocurrencias del experimento. Formalmente, este concepto se define como variable aleatoria y su estructura probabilística es definida a través de las funciones de distribución.

Generalmente, las variables aleatorias son representadas por letras mayúsculas de la parte final del abecedario, tales como X, Y ó Z . A diferencia de las variables (no aleatorias) o incógnitas de una ecuación, que generalmente denotamos por letras minúsculas x, y ó z . Sin embargo, es importante comprender que una variable aleatoria es, en realidad, una función del espacio muestral Ω en los números reales \mathbb{R} .

Definición 1.3 (Variable Aleatoria)

Considere el espacio de probabilidad $(\Omega, \mathcal{F}, \mathcal{P})$. Llamamos *variable aleatoria* a cualquier función $X: \Omega \rightarrow \mathbb{R}$, tal que para todo intervalo $I \subset \mathbb{R}$, se cumple que

$$X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\} \in \mathcal{F}. \quad (1.10)$$

Es decir, X es una función definida de tal forma que para todo intervalo $I \subset \mathbb{R}$, su pre-imagen pertenece a la σ -álgebra \mathcal{F} de Ω .

Note que la función X , en general, es no biyectiva, por lo que su función inversa no existe. En efecto, si una variable aleatoria es definida en un conjunto finito $A \subset \mathbb{R}$, por ejemplo, los valores $A = \{0,1\}$ (variable Bernoulli), el complemento A^c tendrá como pre-imagen el conjunto vacío. Recuerde que basados en las propiedades (a) y (b) de la Definición 1.1, sabemos que $\emptyset \in \mathcal{F}$. Por lo tanto, destacamos que la notación en (1.10) es de carácter conjuntista al denotar una pré-imagen. Garantizamos el cálculo de probabilidades con variables aleatorias al exigir que, para cualquier $I \subset \mathbb{R}$, el conjunto $X^{-1}(I)$ sea un evento, o sea un elemento de la σ -álgebra \mathcal{F} (ver Definición 1.2).

Ejemplo 1.8

Considere el experimento aleatorio del Ejemplo 1.1. Defina las variables aleatorias (V.A's),

X : producto de los valores obtenidos en los dos lanzamientos del dado.

Y : máximo entre los valores obtenidos en los dos lanzamientos del dado.

De esta manera, entendemos que las V.A's X e Y asignan a cada elemento de Ω un número real. En este caso, por ejemplo, $X(\{2,3\}) = 6$ ó $X(\{5,2\}) = 10$. Se deduce que $1 \leq X(\omega) \leq 36$, para todo $\omega \in \Omega$.

Así también, $Y(\{4,1\}) = 4$ ó $Y(\{5,6\}) = 6$. Esto es, $Y(\omega) \in \{1,2,3,4,5,6\}$.

Ejemplo 1.9

Considere el EA del Ejemplo 1.5. De forma análoga, interprete el experimento aleatorio de rendir la prueba en el curso B. Considere el espacio muestral en (1.4), defina las V.A's,

X : nota del tercer examen en el curso A, (1.11)

Y : nota del tercer examen en el curso B. (1.12)

Observe que ambas variables aleatorias son definidas mediante la función identidad. Esto es, por ejemplo, $X(\omega) = \omega$. Esto porque en ambos casos, tanto el espacio muestral como la V.A representan el mismo objeto. De forma similar, podemos asociar una variable aleatoria para el espacio muestral del Ejemplo 1.4.

En este sentido, en muchas situaciones se evita la necesidad de usar los conceptos de experimento aleatorio, espacio muestral y la definición formal de la función $X(\omega)$. Esto es, se define directamente una variable aleatoria X sin especificar Ω , el cual debe ser interpretado de la manera trivial, como en el ejemplo anterior. A manera de ejemplo, mencionamos las siguientes variables aleatorias:

X_1 : número de goles de un equipo de futbol en una temporada,

X_2 : número de tweets publicados desde la ciudad de Temuco en un día, (1.13)

X_3 : tiempo entre los arribos (consecutivos) de dos buses a un paradero.

Por tanto, como de manera usual, omitiremos la mención a ω en el evento $\{\omega \in \Omega : X(\omega) \in I\}$, siendo denotado por $[X \in I]$. Por ejemplo, cuando denotamos $\mathcal{P}(X < 0)$, simplificamos la notación $\mathcal{P}(\{\omega \in \Omega : X(\omega) \in (-\infty, 0)\})$.

Finalmente, advertimos que el cálculo de las probabilidades con variables aleatorias dependerá de la naturaleza de la variable en estudio. Por esto, necesitamos clasificar las variables aleatorias en dos tipos, esto es, separar los casos discreto y continuo.

Definición 1.4 (Variable aleatoria discreta y función de probabilidad)

Una variable aleatoria será definida como *discreta*, si solamente asume un número contable de valores (finito o infinito). La *función de probabilidad*, $p(\cdot)$, de una variable aleatoria discreta es una función que atribuye probabilidad a cada uno de los posibles valores asumidos por la variable. Esto es, sea X una variable con valores x_1, x_2, \dots , tenemos para $i = 1, 2, \dots$,

$$p(x_i) = \mathcal{P}(X = x_i) = \mathcal{P}(\{\omega \in \Omega : X(\omega) = x_i\}). \quad (1.14)$$

Antes de definir el caso continuo, precisamos introducir el concepto de función de distribución. Conocer la función de distribución implica obtener cualquier información sobre la variable aleatoria. En este sentido, como veremos, la variable aleatoria continua será definida mediante su función de distribución.

Definición 1.5 (Función de Distribución Acumulada)

Sea X una variable aleatoria en $(\Omega, \mathcal{F}, \mathcal{P})$, su función de distribución acumulada, o simplemente *función de distribución* es dada por

$$F_X(x) = \mathcal{P}(X \in (-\infty, x]) = \mathcal{P}(X \leq x), \quad (1.15)$$

con x recorriendo los números reales.

Aunque en algunos casos la variable aleatoria X solo asume valores en un subconjunto de los reales, la función de distribución es definida en toda la recta. En el caso de una variable aleatoria discreta, su función de distribución es dada por

$$F_X(x) = \sum_{x_i \leq x} p(x_i),$$

donde $p(\cdot)$ es la función de probabilidad definida en (1.14). A seguir, el caso continuo.

Definición 1.6 (Variable aleatoria continua y función de densidad)

Una variable aleatoria X es *continua* si su recorrido es un conjunto no numerable. Decimos que el conjunto de posibles valores de la variable abarca todo un intervalo de números reales. Formalmente, sea F_X su función de distribución, X es continua si existe una función no negativa f tal que:

$$F_X(x) = \int_{-\infty}^x f(w)dw, \quad (1.16)$$

para todo $x \in \mathbb{R}$. La función f es denominada *función de densidad*.

Como concepto final, incluimos la definición de un elemento esencial en las áreas de probabilidad y estadística, esto es, la noción de esperanza matemática.

Definición 1.7 (Esperanza de una variable aleatoria)

- Caso discreto: Sea X una V.A discreta con función de probabilidad p_X y valores x_i , para i perteneciente a un conjunto de índices \mathcal{I} . El valor esperado, media o esperanza matemática de X es dado por

$$E(X) = \sum_{i \in \mathcal{I}} x_i p_X(x_i), \tag{1.17}$$

siempre que la suma exista ($E(X) < \infty$).

- Caso continuo: Sea X una V.A continua con función de densidad $f(x)$. Definimos el valor esperado de X por

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx, \tag{1.18}$$

siempre que la integral esté bien definida.

Esta cantidad es frecuentemente utilizada como resumen del comportamiento de la variable aleatoria y, como veremos en la Sección 1.7, sirve como parámetro para construir los modelos de probabilidad. Es interesante notar que la esperanza matemática de una V.A, tanto discreta como continua, puede ser asociada con el concepto físico de centro de masa.

1.4 La noción de independencia en probabilidad

Para finalizar esta sección, definimos el concepto de independencia, el cuál posee gran importancia tanto práctica como teórica. En primer lugar, nos referimos a la independencia entre eventos.

Definición 1.8

Los eventos A y B en $(\Omega, \mathcal{F}, \mathcal{P})$ son *independientes* si la información de la ocurrencia de B no altera la probabilidad atribuída al evento A . Esto es,

$$\mathcal{P}(A|B) = \mathcal{P}(A). \tag{1.19}$$

En esta definición usamos el concepto de probabilidad condicional $\mathcal{P}(A|B)$, que representa la probabilidad de ocurrir A dado que el evento B ocurrió. Cuando $\mathcal{P}(B) > 0$, entonces:

$$\mathcal{P}(A|B) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)}. \tag{1.20}$$

Dando interpretación al concepto de independencia entre eventos. Si existe independencia del tipo (1.19), la información que contiene la ocurrencia de B , no afecta la probabilidad atribuída a A anteriormente.

Usando (1.20), la condición en (1.19) puede escribirse de la siguiente manera

$$\mathcal{P}(A \cap B) = \mathcal{P}(A)\mathcal{P}(B). \tag{1.21}$$

A seguir, precisamos definir el concepto de independencia de variables aleatorias. En general, este supuesto es ampliamente usado en las teorías de probabilidad e inferencia estadística. Asumir que las variables aleatorias X e Y son independientes quiere decir que la ocurrencia de (o información sobre) una variable no afecta la distribución de probabilidad de la otra.

Definición 1.9 (Independencia entre variables)

Dos variables aleatorias X e Y son independientes si y solo si (ssi) para cada $a, b \in \mathbb{R}$, los eventos $\{\omega \in \Omega : X(\omega) \leq a\}$ y $\{\omega \in \Omega : Y(\omega) \leq b\}$ son independientes (como en la Definición 1.8). Esto es,

$$\mathcal{P}(X \leq a | Y \leq b) = \mathcal{P}(X \leq a). \quad (1.22)$$

De otra forma, si X e Y poseen funciones de distribución $F_X(x)$ y $F_Y(y)$ respectivamente, son dichas independientes ssi el vector aleatorio (X, Y) posee función de distribución conjunta

$$F_{X,Y}(x, y) = F_X(x)F_Y(y), \quad (1.23)$$

para todo $x, y \in \mathbb{R}$.

Equivalentemente, para el caso discreto usamos las funciones de probabilidad. Entonces, las variables aleatorias discretas X e Y son independientes ssi

$$p_{X,Y}(x, y) = p_X(x)p_Y(y), \quad (1.24)$$

para todo $x, y \in \mathbb{R}$. En el caso continuo y considerando las funciones de densidad. Las variables aleatorias continuas X e Y son independientes ssi, para todo $x, y \in \mathbb{R}$,

$$f_{X,Y}(x, y) = f_X(x)f_Y(y). \quad (1.25)$$

Para concluir esta sección, recordamos que los conceptos hasta ahora definidos, serán utilizados para entender la importancia de los supuestos en la creación y resolución de problemas probabilísticos. Sin olvidar que de forma general, estos contenidos son la base de la teoría matemática de la probabilidad.

1.5 Suposiciones de independencia y de distribución idéntica

En esta sección asumimos el supuesto de que las variables aleatorias X e Y (ver (1.11) y (1.12)) son independientes e idénticamente distribuidas (denotado por i.i.d). Este requisito tiende a simplificar la matemática subyacente.

Destacamos que el requisito de i.i.d es comunmente usado en la teoría frecuentista de inferencia estadística. En general, obtenemos una muestra de X_1, X_2, \dots, X_n , esto es, una secuencia de n realizaciones de una variable aleatoria. En este sentido, es posible suponer que cada realización (del mismo experimento aleatorio) es independiente de las otras y como observamos la misma variable, la distribución de probabilidad debe ser la misma para cada observación. Bajo este supuesto existen dos resultados de convergencia (n tiende a ∞) muy importantes en la teoría de probabilidad y la estadística inferencial, esto es, la *ley de los grandes números* y el *teorema central del límite*.

En el caso del problema que nos motiva, el supuesto de que X e Y son i.i.d fue implícitamente asumido en Malaspina [19]. Considerar independencia entre las variables aleatorias X e Y parece ser razonable. O sea, decimos que la nota obtenida en el curso A no interfiere en la probabilidad de la nota obtenida en el curso B (y viceversa). No obstante, el supuesto de que ambas variables tienen la misma función de distribución es cuestionable.

En el trabajo citado, de forma poco rigurosa, se utiliza un par de modelos de probabilidad (a ser definido en la Sección 1.7) para responder a la interrogante formulada en la Sección 1.2. Sin embargo, ahora veremos que bajo la suposición de que las variables aleatorias X e Y son independientes e idénticamente distribuidas, la pregunta posee una respuesta trivial.

De otra forma, fácilmente podemos garantizar que:

El curso B tiene mayor probabilidad de ser aprobado.

Resolución: Asumiendo que X e Y son i.i.d, estamos diciendo que $F_X = F_Y$, como definidos en (1.15). De otra forma, denotando \mathcal{P}_X y \mathcal{P}_Y sus respectivas distribuciones de probabilidad asociadas, decimos que $\mathcal{P}_X = \mathcal{P}_Y := \mathcal{P}$. Una vez que para aprobar A, necesitamos $\{X \geq 10.85\}$ y para aprobar B, se necesita $\{Y \geq 8.85\}$ (ver (1.2) y (1.3)). Claramente, aplicando la propiedad (P3) de la Proposición 1.1 y luego el supuesto de distribución idéntica, obtenemos

$$\mathcal{P}(X \geq 10.85) \leq \mathcal{P}(X \geq 8.85) = \mathcal{P}(Y \geq 8.85). \quad (1.26)$$

Con esto, debemos concluir que la interrogante planteada en la Sección 1.2 se torna trivial, una vez asumidos los supuestos de independencia y de distribuciones idénticas.

Parece razonable la opción de eliminar la suposición de que las distribuciones son idénticas, pero como hemos dicho, incluso este camino nos llevará a deducir que el curso B tiene mayor probabilidad de ser aprobado. Esto por la tendencia de obtener notas más altas en dicho curso, lo que sugiere usar una distribución asimétrica para la variable aleatoria Y . Esta idea será discutida en la Sección 1.8.

El camino es abandonar esta pregunta. A seguir, veremos que basados en el enunciado de la Sección 1.2 (y propuesto en Malaspina [19]), es posible formular preguntas simples que nos llevan a desarrollos matemáticos interesantes para tratar diversos contenidos de la teoría de probabilidad.

1.6 El desafío de la creación de problemas de probabilidad

En la Sección 1.3 formulamos las bases de la teoría de probabilidad, esto es, la construcción axiomática del concepto de probabilidad, la definición de variable aleatoria y su función de distribución. De forma complementaria estudiamos la noción de independencia probabilística.

Los fundamentos matemáticos de la base teórica de la probabilidad, deben entenderse como conceptos elementales de dicha área. Estos fundamentos requieren el conocimiento de varias áreas de la matemática, tales como la teoría de conjuntos y la teoría de la medida. Por otra parte, veremos que

en sus aplicaciones serán necesarios conceptos frecuentemente estudiados en los cursos de cálculo. Además de los prerrequisitos en teoría combinatoria como requiere la probabilidad clásica de Laplace. El conocimiento de estos contenidos, y cierto grado de dominio, son esenciales para comprender el nivel de dificultad que implica la creación de problemas de probabilidad.

En la segunda parte de nuestro trabajo (Sección 1.5), vimos como los supuestos considerados en el planteamiento de un problema de probabilidad, nos pueden llevar a preguntas cuya solución es trivial. Entendiendo que el camino de creación de problemas presenta diversas opciones, una de ellas es la posibilidad de incluir o abandonar ciertos supuestos para crear problemas que se tornen considerablemente más complejos. Enfrentar esa complejidad es un desafío que permite al docente explorar diversos conceptos matemáticos. De esta forma, consideramos que volviendo más realista la situación, complejizamos la matemática.

En esta sección, presentamos dos interrogantes que derivan del enunciado en la Sección 1.2. Ambas preguntas son razonables de ser propuestas por estudiantes. Nuestro objetivo es comprender como los conceptos de variable aleatoria y distribución de probabilidad son utilizados para aplicar dicha teoría. Para tal motivo, se incluye la definición y propiedades de algunos modelos de probabilidad usualmente referidos en la bibliografía matemática.

A seguir, presentamos nuestra propuesta de problemas a ser estudiados. Ambas interrogantes serán respondidas en la Sección 1.8.

(1) *¿Qué es más probable, aprobar el curso A o reprobado el curso B?*

(2) *Suponga que una beca es mantenida si el promedio, sin redondeo, de las 6 notas (3 de cada curso) es superior a 13. ¿Cuál es la probabilidad de que el alumno mantenga la beca?*

El problema (1) es una forma de complementar el estudio de la sección anterior. En el caso del problema (2) buscamos entender los conceptos de distribuciones conjuntas e independencia de variables aleatorias. Para resolver estas interrogantes consideramos el supuesto de que las distribuciones no son idénticas. En este sentido, presentamos algunas distribuciones de probabilidad frecuentemente mencionadas en la literatura matemática.

1.7 Modelando la aleatoriedad

Una vez que hemos definido las variables aleatorias X e Y , representando las notas de los cursos A y B respectivamente, necesitamos precisar sus funciones de distribución F_X y F_Y . De otra forma, si la variable aleatoria es discreta (o continua) debemos conocer la función de probabilidad (o función de densidad), como en la Definición 1.4 (o Definición 1.6).

En general, sea el caso discreto o continuo, una especificación de dicha función será llamada modelo o distribución de probabilidad. En este apartado presentamos algunas distribuciones de probabilidad usualmente aplicadas.

Antes de esto, definiremos dos funciones matemáticas que serán utilizadas en la secuencia. La primera de ellas es llamada *función indicadora* de un conjunto A y es dada por:

$$I_A(\omega) = \begin{cases} 1, & \text{si } \omega \in A, \\ 0, & \text{si } \omega \notin A. \end{cases} \tag{1.27}$$

La siguiente función es llamada *función gamma* y es definida para todo $t > 0$,

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx. \tag{1.28}$$

Usando integración por partes es fácil ver que $\Gamma(t + 1) = t \cdot \Gamma(t)$. De esta forma, si t es un número entero positivo obtenemos,

$$\Gamma(t + 1) = t \times (t - 1) \times (t - 2) \times \dots \times 1 = t!. \tag{1.29}$$

Estamos en condiciones de definir los modelos de probabilidad. El primer caso que incluimos es la distribución uniforme discreta. Este modelo representa situaciones en que todos los posibles valores de la variable aleatoria son *equiprobables*. Es una distribución simple, que puede ser comparada con la noción de *probabilidad clásica* (1.6) dada por Laplace.

Definición 1.10 (Distribución uniforme discreta)

Una variable Z sigue una distribución uniforme discreta, con valores z_1, z_2, \dots, z_k , si tiene función de probabilidad dada por:

$$p(z_i) = \frac{1}{k} I_E(i), \tag{1.30}$$

donde $E = \{1, 2, \dots, k\}$. Usamos la notación $Z \sim Unif(E)$, con E siendo el conjunto de sus posibles valores.

Observe que este modelo es asumido en el típico ejemplo de lanzamiento de un dado, con $z_i = i \in \{1, 2, 3, 4, 5, 6\}$.

En la secuencia, presentaremos dos distribuciones continuas. La primera de ellas es la distribución Normal, con importancia inmensurable para la estadística y la ciencia en general. La segunda distribución es llamada Beta y su flexibilidad para ajustar diferentes curvas de datos la tornan muy útil en diversas áreas.

Definición 1.11 (Distribución Normal)

Una variable aleatoria Z sigue una distribución Normal si su función de densidad es dada por

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} I_{(-\infty, \infty)}(z), \tag{1.31}$$

donde $\mu, \sigma \in \mathbb{R}, \sigma > 0$. Usaremos la notación $Z \sim N(\mu, \sigma^2)$.

En este caso, el parámetro μ representa la media y σ^2 representa la varianza de la variable Z . Como podemos verificar en la Figura 1.1, la distribución Normal es simétrica.

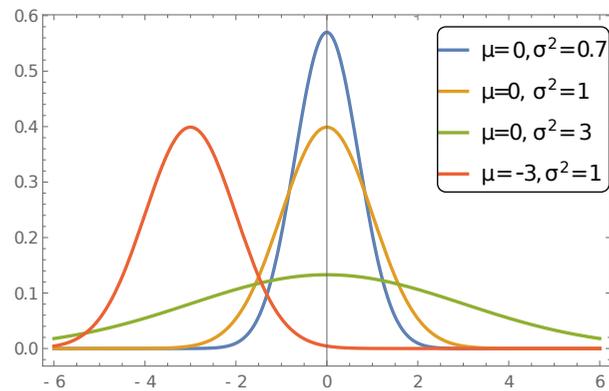


Figura 1.1: Funciones de densidad para diferentes valores de los parámetros. (a) Distribución Normal, $Z \sim N(\mu, \sigma^2)$.

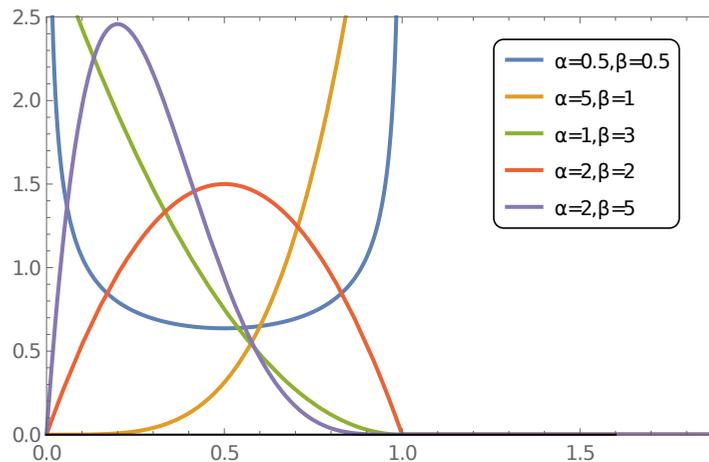


Figura 1.2: Distribución Beta, $Z \sim Beta(\alpha, \beta)$.

Definición 1.12 (Distribución Beta)

Una variable aleatoria Z tiene distribución Beta, si su función de densidad es de la forma,

$$f(z) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} z^{\alpha-1}(1-z)^{\beta-1} I_{[0,1]}(z), \quad (1.32)$$

donde los parámetros α y β son números reales positivos. La notación usada es $Z \sim Beta(\alpha, \beta)$. La esperanza es dada por $E(Z) = \alpha / (\alpha + \beta)$.

Como hemos dicho, la distribución Beta presenta una gran flexibilidad para ajustar curvas con diversas formas (ver Figura 1.2). Dependiendo de los valores de los parámetros α y β podemos obtener funciones de densidad simétricas o asimétricas, tanto cóncavas como convexas.

Debemos mencionar que, a pesar de la trivialidad verificada en la Sección 1.5, en Malaspina [19] se presenta una solución a la pregunta original bajo la suposición de normalidad, es fácil ver que el supuesto es poco razonable dado el espacio de valores que asume una V.A con distribución Normal, esto es, los números reales \mathbb{R} . Además, se está restringiendo a la suposición de que las notas son simétricas en torno a la media.

En nuestro estudio, la distribución Beta será la base para modelar las notas del tercer examen en los cursos A y B , esto es, las variables aleatorias X e Y , respectivamente. Para dicho fin usamos el concepto de transformación de una V.A. En este sentido, observe que $Z \sim \text{Beta}(\alpha, \beta)$ asume valores en el intervalo $[0, 1]$, pero podemos sugerir una transformación simple de la variable aleatoria. Esto es, usaremos $X = 20 \cdot Z$, que asumirá valores en el intervalo $[0, 20]$. En general podemos utilizarlo para cualquier intervalo $[0, a]$, con $a > 0$.

Por otra parte, cerramos esta sección comentando el interés que poseen las distribuciones de Poisson (discreta) y exponencial (continua) en el área de procesos estocásticos (ver, por ejemplo, [18, 30, 31]). La importancia se debe a sus innumerables utilidades en las áreas de probabilidad aplicada e ingeniería. Por ejemplo, las llamadas telefónicas entrantes en una central o la llegada de clientes a la fila de un banco, pueden ser modeladas satisfactoriamente a través de los *procesos de Poisson* (ver, Alabert [1]). Un proceso de Poisson posee la propiedad de que el número de llegadas sigue una distribución de Poisson (vea las variables X_1 y X_2 en (1.13)) y, el tiempo de espera entre dos llegadas (consecutivas) sigue una distribución exponencial (ver X_3 en (1.13)).

1.8 Resolución de problemas y los supuestos en probabilidad

En esta sección utilizaremos las herramientas matemáticas definidas en las secciones anteriores, para dar respuesta a las interrogantes (1) y (2) enunciadas en la Sección 1.6.

Primero suponemos que las notas serán modeladas por valores reales, o sea, usamos el espacio muestral Ω del Ejemplo 1.5. Así, X e Y son variables aleatorias continuas. También usamos el supuesto de independencia entre las V.A.'s. Sin embargo, suponemos que las distribuciones de probabilidad \mathcal{P}_X y \mathcal{P}_Y , asociadas a cada variable son diferentes.

Así, asumimos que $X = 20 \cdot Z_A$ e $Y = 20 \cdot Z_B$, donde Z_A, Z_B siguen distribuciones Beta con diferentes parámetros. De forma arbitraria, y a modo de ejemplo, basados en las notas de los exámenes anteriores utilizaremos los siguientes supuestos distribucionales¹ (el lector puede considerar otros parámetros).

$$Z_A \sim \text{Beta}(2,2) \quad \text{y} \quad Z_B \sim \text{Beta}(3,2), \quad (1.33)$$

ver la Figura 1.3. Denotamos por f_{Z_A} y f_{Z_B} , sus respectivas funciones de densidad (ver (1.32)).

De esta forma, incluimos una asimetría asociada a la tendencia de obtener mejores notas en el curso B . Basados en las medias aritméticas de las notas anteriores (ver Tabla 1.1), hemos asumido los valores de los parámetros en (1.33) para que $E(X) = 10$ y $E(Y) = 12$.

¹Desde el punto de vista de la teoría estadística, la escasa cantidad de datos impide que podamos fundamentar nuestra elección. Sin embargo, el motivo de este trabajo es comprender como se construye la matemática una vez establecidas las suposiciones. Además nuestro foco es modelar variables aleatorias y no ajustar distribuciones a datos empíricos. Con esto, justificamos dicha arbitrariedad.

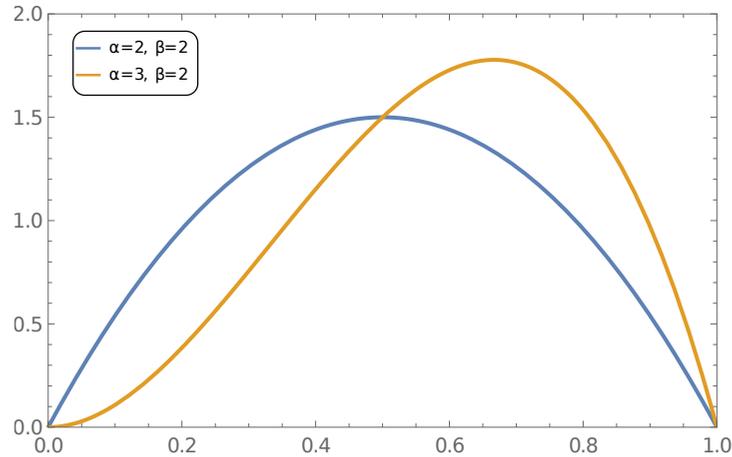


Figura 1.3: Funciones de densidad de las variables aleatorias Z_A (en celeste) y Z_B (en amarillo).

1.8.1 Respuesta a la pregunta (1)

Para responder a la primera pregunta debemos considerar las inecuaciones (1.2) y (1.3). De otra forma, necesitamos calcular las siguientes probabilidades.

$$\mathcal{P}_X(X \geq 12.85) = \mathcal{P}\left(Z_A \geq \frac{12.85}{20}\right), \quad (1.34)$$

denotando la probabilidad de aprobar el curso A. Además, la probabilidad de reprobar el curso B,

$$\mathcal{P}_Y(Y < 8.85) = \mathcal{P}\left(Z_B < \frac{8.85}{20}\right). \quad (1.35)$$

Una vez que hemos supuesto que X e Y son variables aleatorias continuas, y sus leyes de probabilidad son asociadas a variables Z_A y Z_B con distribución Beta, debemos usar la Definición 1.5, Definición 1.6 y la función de densidad (1.32) para calcular estas probabilidades. En este sentido, usando (1.29) y resolviendo la integral de un polinomio,

$$\begin{aligned} \mathcal{P}(Z_A \leq a) &= \int_0^a \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)} z(1-z) dz \\ &= 6 \left(\frac{z^2}{2} - \frac{z^3}{3} \right) \Big|_0^a \\ &= 6a^2 \left(\frac{1}{2} - \frac{a}{3} \right), \end{aligned} \quad (1.36)$$

para todo $a \in [0, 1]$. Observe que, en concordancia con los axiomas de la Definición 1.2, $\mathcal{P}(Z_A \leq 1) = 1$. De forma análoga, para todo $b \in [0, 1]$

$$\mathcal{P}(Z_B < b) = \int_0^b \frac{\Gamma(5)}{\Gamma(3)\Gamma(2)} z^2(1-z) dz = 12 \left(\frac{z^3}{3} - \frac{z^4}{4} \right) \Big|_0^b = 12b^3 \left(\frac{1}{3} - \frac{b}{4} \right). \quad (1.37)$$

Finalmente, usamos (1.36) para obtener

$$\mathcal{P}_X(X \geq 12.85) = 1 - \mathcal{P}\left(Z_A \leq \frac{12.85}{20}\right) = 1 - 6 \left(\frac{12.85}{20} \right)^2 \left(\frac{1}{2} - \frac{12.85}{60} \right) = 0.292\dots \quad (1.38)$$

Luego usamos (1.37),

$$\mathcal{P}_Y(Y < 8.85) = \mathcal{P}\left(Z_B < \frac{8.85}{20}\right) = 0.231\dots \quad (1.39)$$

La conclusión es clara.

Es más probable aprobar el curso A que reprobar el curso B.

Observación: A modo de ejercicio, usando (1.36) y (1.37), el lector puede responder a la pregunta original (en la Sección 1.2) asumiendo nuestros supuestos distribucionales. Claramente, la conclusión será la misma que expusimos en la Sección 1.5. Por otra parte, es posible ejercitar respondiendo a nuestra Pregunta (1) asumiendo que X e Y son variables aleatorias i.i.d.

1.8.2 Respuesta a la pregunta (2)

Para mantener la beca, el alumno necesita $\frac{1}{6}(8 + 12 + X + 10 + 14 + Y) \geq 13$. O sea, $X + Y \geq 34$. Nuestra labor es calcular

$$\mathcal{P}(X + Y \geq 34) = \mathcal{P}(Z_A + Z_B \geq 1.7). \tag{1.40}$$

Dada la independencia asumida entre X e Y , las variables Z_A y Z_B son también independientes. Denotamos f_{Z_A, Z_B} la función de densidad conjunta (ver (1.25)) del vector aleatorio (Z_A, Z_B) . De esta forma,

$$\mathcal{P}(Z_A + Z_B \geq 1.7) = \iint_{\{z+w \geq 1.7\}} f_{Z_A, Z_B}(z, w) \, dzdw = \int_{0.7}^1 \int_{(1.7-z)}^1 f_{Z_A}(z) f_{Z_B}(w) \, dw dz. \tag{1.41}$$

Por lo tanto, necesitamos calcular una integral doble para responder a la interrogante. Usaremos algunas propiedades de la probabilidad para reducir los cálculos. Entonces,

$$\begin{aligned} \int_{0.7}^1 \int_{(1.7-z)}^1 f_{Z_A}(z) f_{Z_B}(w) \, dw dz &= \int_{0.7}^1 f_{Z_A}(z) \left[\int_{(1.7-z)}^1 f_{Z_B}(w) \, dw \right] dz \\ &= \int_{0.7}^1 f_{Z_A}(z) (1 - \mathcal{P}(Z_B < 1.7 - z)) dz \\ &= \int_{0.7}^1 f_{Z_A}(z) dz - \int_{0.7}^1 f_{Z_A}(z) \mathcal{P}(Z_B < 1.7 - z) dz \\ &= (1 - \mathcal{P}(Z_A \leq 0.7)) - \int_{0.7}^1 f_{Z_A}(z) \mathcal{P}(Z_B < 1.7 - z) dz \end{aligned} \tag{1.42}$$

Finalmente, podemos utilizar los resultados de (1.36) y (1.37). Así, juntando (1.41) y (1.42), decimos que

$$\mathcal{P}(Z_A + Z_B \geq 1.7) = \left(1 - 6(0.7)^2 \left(\frac{1}{2} - \frac{0.7}{3} \right) \right) - \int_{0.7}^1 6z(1-z) \cdot 12 \frac{(1.7-z)^3}{3} dz + \int_{0.7}^1 6z(1-z) \cdot 12 \frac{(1.7-z)^4}{4} dz$$

Un cambio de variable y resolver una integral de un polinomio nos lleva a concluir que,

$$\mathcal{P}(Z_A + Z_B \geq 1.7) = \mathcal{P}(X + Y \geq 34) = 0.017. \tag{1.43}$$

Bajo el requisito de independencia y los supuestos distribucionales que hemos asumido, podemos decir que la probabilidad de que el alumno mantenga su beca es baja, igual a 0.017. En otras palabras, las chances de mantener la beca no alcanzan un 2%.

1.9 Conclusiones y comentarios

En este trabajo hemos presentado un enfoque para la enseñanza de la probabilidad basados en la resolución de problemas. Particularmente, destacamos la importancia de los supuestos a la hora de crear y resolver problemas en el área de probabilidades. Como vimos, las suposiciones asumidas para resolver un problema, pueden tornarlo una trivialidad. Así como también, otro tipo de supuestos permiten plantear diversos problemas matemáticos que requieren aplicar los diferentes contenidos de dicha área. Sin embargo, es importante destacar que las suposiciones que deben ser asumidas, serán aquellas que se adecuen de mejor manera a la realidad, implicando o no, la necesidad de herramientas matemáticas complejas.

De otra forma, consideramos como motivación un problema de probabilidad estudiado en Malaspina [19]. En dicho artículo son discutidas diversas soluciones, que se resumen a considerar los casos discreto y continuo para el espacio muestral en el Ejemplo 1.5 (como hicimos con Ω' y Ω). Es preciso comprender que en [19] siempre se enfrenta el problema asumiendo los supuestos de independencia y de distribución idéntica (i.i.d). Con esto, la respuesta siempre es la misma como hemos vistos. En la secuencia propusimos un par de nuevas interrogantes que fueron resueltas asumiendo ciertos supuestos distribucionales para obtener soluciones adecuadas a la información disponible.

En la primera parte del texto expusimos una revisión de los conceptos básicos de la teoría de probabilidad. Referencias para profundizar la matemática de dicha teoría son los libros de L. Rincón [27], B. James [9] y S. Ross [30].

Notamos que los fundamentos dados por Kolmogorov fueron cruciales para el desenvolvimiento de la probabilidad. Los axiomas de la Definición 1.2 han sido ampliamente aceptados, pero existen formulaciones alternativas. En particular, la teoría subjetivista fundada por Bruno de Finetti [6], una discusión filosófica-matemática de dicha teoría se encuentra en el libro de Cofré y Henao [4].

En la Sección 1.5, nos referimos a las suposiciones de independencia y distribuciones idénticas, mencionando dos importantes resultados de la teoría de probabilidad: la ley de los grandes números (LGN) y el teorema central del límite (TCL). En este sentido, si consideramos una secuencia i.i.d. de variables aleatorias X_1, X_2, \dots, X_n , con esperanza μ y varianza σ^2 , definimos $S_n = X_1 + \dots + X_n$. La (versión débil de la) LGN dice que, si tomamos $n \rightarrow \infty$, la probabilidad de que el valor de $n^{-1}S_n$ se concentre en una vecindad de μ tiende a 1. Este resultado muestra una forma de equilibrio asintótico en torno a la media. En el caso del TCL, garantizamos que la distribución de $\sqrt{n}(S_n - \mu)$ converge a una distribución Normal con media cero y varianza σ^2 , sin importar la distribución que asumamos para las variables X_i que conforman la secuencia i.i.d.

Como complemento a la Sección 1.6, plantemos el desafío a los estudiantes de pedagogía para simular computacionalmente algunas variables aleatorias, con la posibilidad de explorar las potencialidades del software estadístico R [25]. Con el cual, por ejemplo, hemos obtenido las funciones de densidad en nuestras figuras. Para los motivos de simulación, citamos el Capítulo 5 de Montes [21], el Capítulo 10 de Ross [30] y el libro de introducción a R [26].

Finalmente, con respecto a la Sección 1.8, decimos que una vez que comenzamos a modelar la aleatoriedad debe resultar claro que las soluciones presentadas en dicha sección dependen fuertemente de los supuestos distribucionales asumidos, específicamente de la elección de los parámetros en (1.33). Si los parámetros escogidos fuesen otros, obtendríamos diferentes valores para las probabilidades calculadas.

De otra forma, la teoría de probabilidad nos ofrece distintos modelos con las más diversas aplicaciones. Las respuestas obtenidas en un cierto estudio dependen directamente de la elección de un modelo específico. El ajuste adecuado de un modelo al problema en estudio es claramente una labor importante. Al momento de escoger arbitrariamente los valores de los parámetros estamos esquivando dicha labor. Este objetivo es el foco principal del área de *inferencia estadística*, esto es, la estimación de parámetros o de funciones de distribución basados en los datos empíricos.

Agradecimientos. El autor agradece a la Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP (Processo 15/02801-6). Además, agradece los comentarios y sugerencias de Natalia Milla (IME/USP, Brasil) y de Eduardo Infante (Colegio Playground, Chile). Finalmente destaca el valioso aporte de los árbitros encargados de la revisión del manuscrito.

Bibliografía

- [1] Alabert, A. "The Modelling of Random Phenomena". *Materials Matemàtics, Revista Electrónica de Divulgación Matemática*, (2015).
- [2] Alsina, A., Vásquez, C. "Análisis de los conocimientos probabilísticos del profesorado de Educación Primaria". *Revista digital Matemática, Educación e Internet*, (2016) 16(1).
- [3] Batanero, C., Ortiz, J.J., Serrano, L. "Investigación en didáctica de la probabilidad". *UNO*, (2007) 44, 7-16.
- [4] Cofré, Á., Henao, D. *Previsión y probabilidad, las ideas de Bruno de Finetti*. Ediciones PUC, Chile. 2007.
- [5] Debnath, L., Basu, K. "A short history of probability theory and its applications". *Int. Journal of Mathematical Education in Science and Technology*, (2015) 46(1), 13-39.
- [6] de Finetti, B. *Theory of Probability*. 2 volúmenes, New York: Wiley. 1974-5.
- [7] Francisco, J.M., Maher, C.A. "Conditions for promoting reasoning in problem solving: Insights from a longitudinal study". *Journal of Mathematical Behaviour*, (2005) 24, 361-372.
- [8] González Navarrete, M. "Propuestas para la enseñanza de las probabilidades: Un ejemplo basado en la educación media chilena". *UNIÓN, Revista Iberoamericana de Educ. Matemática*, (2013) 36, 145-163.
- [9] James, B. *Probabilidade: um curso de nível intermediário (3a ed.)*. Rio de Janeiro: IMPA, 2006.
- [10] Jiménez, L., Jiménez, J.R. "Enseñar probabilidad en primaria y secundaria?, ¿Para qué y por qué?". *Revista digital Matemática, Educación e Internet*, (2005) 6(1).
- [11] Jiménez Saavedra, N. "La axiomática de Kolmogorov: fundamento de la teoría de la probabilidad. Las Matemáticas del siglo XX". *NÚMEROS, Revista de Didáctica de las Matemáticas*, (2000) 43, 185-190.
- [12] Jones, G. (ed.) *Exploring probability in schools: Challenges for teaching and learning*, 15-37. Springer: New York. 2005.
- [13] Klymchuk, S., Kachapova, F. "Paradoxes and counterexamples in teaching and learning of probability at university". *Int. Journal of Mathematical Education in Science and Technology*, (2012) 43(6), 803-811.
- [14] Kolmogorov, A. *Foundations of the Theory of Probability (2a ed.)*. New York: Chelsea. 1956.

- [15] Laplace, P. *Théorie Analytique des Probabilités*. Courcier, Paris. 1812.
- [16] Laplace, P. *Essai philosophique sur les probabilités*. Courcier, Paris. 1814.
- [17] Lesh, R., Doerr, H.M. (Eds.) *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching*. Mahwah, NJ: Lawrence Erlbaum. 2003.
- [18] Magalhães, M.N. *Probabilidade e variáveis aleatórias*. Edusp, São Paulo. 2006.
- [19] Malaspina, U. "Creación de problemas. Un caso con probabilidades". Sección: El rincón de los problemas. *UNIÓN, Revista Iberoamericana de Educ. Matemática*, (2013) 33, 119-124.
- [20] Malaspina, U. "La creación de problemas de matemáticas en la formación de profesores". VII Congreso Iberoamericano de Educación Matemática. Montevideo, (2013), 117-128.
- [21] Montes, F. *Introducción a la probabilidad*. Universitat de València. 2007. Disponible en <http://www.uv.es/montes/probabilitat/manual.pdf>
- [22] Nunokawa, K. "Mathematical problem solving and learning mathematics: What we expect students to obtain". *Journal of Mathematical Behavior*, (2005) 24(3-4), 325-340.
- [23] Ollerton, R. "A unifying framework for teaching probability event types". *Int. Journal of Mathematical Education in Science and Technology*, (2015) 46(5), 790-794.
- [24] O'Shea, J., Leavy, A.M. "Teaching mathematical problem-solving from an emergent constructivist perspective: the experiences of Irish primary teachers". *Journal of Mathematics Teacher Education*, (2013) 16(4), 293-318.
- [25] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2015. <http://www.R-project.org/>
- [26] R Development Core Team. *Introducción a R*. (Versión 1.0.1) 2000. Disponible en <https://cran.r-project.org/doc/contrib/R-intro-1.1.0-espanol.1.pdf>
- [27] Rincón, L. *Curso intermedio de probabilidad*. Las Prensas de Ciencias (UNAM), México, D.F. 2015. Disponible en <http://lya.fciencias.unam.mx/lars/Publicaciones/cip2010.pdf>
- [28] Rincón, L. *Introducción a la probabilidad*. Las Prensas de Ciencias (UNAM), México, D.F. 2015. Disponible en <http://lya.fciencias.unam.mx/lars/Publicaciones/Prob1-2014.pdf>
- [29] Rincón, L. *Introducción a los procesos estocásticos*. Las Prensas de Ciencias (UNAM), México, D.F. 2012. Disponible en <http://lya.fciencias.unam.mx/lars/Publicaciones/procesos2012.pdf>
- [30] Ross, S. *A First Course in Probability (9a ed.)*. Pearson Prentice Hall, New Jersey. 2013.
- [31] Ross, S. *Introduction to Probability Models (10a ed.)*. Academic Press, Inc. Orlando. 2009.
- [32] Santaló, L. *Hacia una didáctica humanista de la matemática*. Troquel S.A, Buenos Aires. 1999.
- [33] Sepúlveda, A., Medina, C., Sepúlveda, D. "La resolución de problemas y el uso de tareas en la enseñanza de las matemáticas". *Educación Matemática*, (2009) 21(2).
- [34] Shafer, G., Vovk, V. *The origins and legacy of Kolmogorov's Grundbegriffe*. 2013. Disponible en <http://www.probabilityandfinance.com/articles/04.pdf>
- [35] Taplin, M., Chan, C. "Developing Problem-Solving Practitioners". *Journal of Mathematics Teacher Education*, (2001) 4(4), 285-304.
- [36] Vargas, C. "Contribuciones de Laplace al desarrollo de la teoría de la probabilidad: en conmemoración de los 200 años de la publicación del *Essai philosophique sur les probabilités* de 1814". *Revista digital Matemática, Educación e Internet*, (2015) 15(2).
- [37] Viali, L., Cury, H.N. "Análise de erros em probabilidade: uma pesquisa com professores em formação continuada". *Educação Matemática Pesquisa*, (2009) 11(2), 373-391.