UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales

Departamento de Matemática

Tesis de Licenciatura

Estudio y diseño de una propuesta didáctica para un abordaje de la Ley de los Grandes Números y el Teorema Central del Límite

Nicolás Igolnikov

Dirección: Dra. Mariela Sued

Agosto 2023

Agradecimientos

Del tesista Esta tesis es el resultado final de un largo trabajo que implicó el apoyo de diversas personas a las cuales estoy considerablemente agradecido. La amistad, el cariño y la confianza han sido fundamentales para poder llevarlo adelante, y su valor no puede atraparse con palabras. Algunas menciones, no obstante, es imposible no hacerlas.
A Carmen Sessa y a Mariela Sued, por alentarme a seguir esta tesis y por el acompañamiento a lo largo de todo el proceso.
A Débora y Mariano, dos pilares sin los cuales nada de todo tendría sentido.
Al hermoso grupo de estudiantes que dimos en llamar Cuerpo Completo[1], y que fue tan importante para tantas cosas.
Al Departamento de Matemática y al CODEP por dar su voto de confianza en este trabajo.
Y a Pola, el agradecimiento infinito.

De la directora

Quiero agradecer a Carmen por haberme acompañado en este proyecto. En particular, por haber supervisado la elaboración de los primeros cuatro problemas incluidos en la secuencia propuesta en el Capítulo 4. En el mismo se presentan construcciones alternativas de algunos conceptos fundacionales (variables binomiales), ofreciendo un recorrido inusual para el abordaje de algunas tematicas clásicas, y se incluye un meticuloso análisis de los mismos. Si bien he acompañado los aspectos matemáticos de la tesis, la colaboración de Carmen fue fundamental, siendo que el enfoque propuesto se ubica mas alla de mi área de trabajo.

Agradezco también a Nico, por todo lo que aprendí a lo largo de los debates que tuvimos. En particular, me enseñó que la tesis es del alumno y no del director.

[1]devenido en Familia no tipo.

Índice general

Capítulo 1

Introducción

El punto de partida de Tesis de Licenciatura es el desarrollo y análisis didáctico-matemático de una propuesta didáctica para el abordaje de la Ley de los Grandes Números y el Teorema Central del Límite en el nivel superior. Pero producir conocimiento que tenga que ver con acciones concretas sobre el sistema educativo (independientemente de la escala o de la ambición de estas acciones) interpela completamente nuestro sistema de conocimientos, desde mucho antes de poder llevarlas a cabo o incluso de concebirlas.

Si le lectore de esta tesis ha sido -o ha tenido la intención de ser- docente en cualquier nivel educativo, sin duda habrá experimentado, al pensarse en posición de estar frente a un aula, el calor de numerosas interrogaciones. ¿Qué sé (o creo saber) en torno a los asuntos que debo enseñar? ¿Qué acciones puedo promover para que otres aprendan? ¿Qué voy a enseñar, y qué relación tendrá ello con lo que les estudiantes aprendan? ¿Cómo aprendí lo que aprendí, cómo otres aprendieron? ¿Qué motivos llevaron a que haya que enseñar esto que hay que enseñar?

Con mayor o menor medida, con mayor o menor marco teórico, con mayor o menos experiencia, todes hemos tenido que sortear la incertidumbre de estas y otras preguntas. Y hemos ido a nuestras aulas así a encontrarnos con la -afortunada- unicidad de cada grupo de estudiantes, de cada clase, de cada curso. Y en estos encuentros, sin duda, hemos aprendido, desaprendido y reaprendido con elles. Esto nos obligó necesariamente a repensar nuestras posiciones, a elaborar nuevas preguntas. Es la sucesión de estos reposicionamientos en parte la que constituye finalmente nuestra manera de actuar en el aula.

En este entramado se encuentra, en un lugar para nada soslayable, el resto del mundo: ¿Cómo responden otres estas preguntas que nosotres nos hacemos? ¿qué hacen otres frente a problemáticas similares a las nuestras? ¿Qué otras preguntas se hacen? ¿Cómo las responderíamos nosotres?

En el caso de este trabajo, nos vemos inmerses en una gran cantidad de producción, a nivel mundial, en torno a la enseñanza y el aprendizaje de la matemática, y en particular de la probabilidad y la estadística. Estos elementos sin duda atraviesan nuestra posición y enmarcan la producción de este trabajo.

Para llevar adelante esta tesis, hemos decidido tomar como punto de partida el sostén de una mirada matemático-didáctica que fuera transversal a todo el proceso, entendiendo que didáctica y matemática no pueden dislocarse a la hora de pensar en la enseñanza-aprendizaje de la matemática. Esta posición constituyó para nosotres un punto de observación y análisis para los diferentes asuntos que se abordan

en esta tesis.
Entre otros múltiples puntos que los diversos desarrollos en el área de la didáctica de la probabilidad y la estadística nos aportaron, destacamos el de la mirada articulada de ambas disciplinas. Pese al lugar propio adquirido por la probabilidad en la ciencia moderna, independiente de la estadística, estas disciplinas se encuentran tanto histórica como actualmente íntimamente relacionadas. A estos fines utilizaremos el término *estocástica*, en consonancia con el amplio uso que este término recibe en la literatura, siempre que los conceptos evocados tengan, al menos para el aula, unos sentidos y maneras de ser vistos desde alguna de estas disciplinas.
Para realizar este trabajo fue necesario realizar un recorte de los asuntos que esta propuesta se propone abordar, que pudiera tomar en cuenta sus (interrelacionadas) dimensiones histórica, matemática, epistemológica y didáctica.
A estos fines, organizamos la tesis en cuatro capítulos:

- Los Problemas
- La matemática y el Aula
- La Propuesta Didáctica
- Las Conclusiones

En el primero, nos proponemos recortar cómo dos problemas de naturaleza estocástica incidieron en el desarrollo de la probabilidad y estadística, y en particular con relación a la formulación de la Ley de los Grandes Números y del Teorema Central del Límite. Este breve recorrido por 140 años de historia matemática nos dará elementos para preguntarnos por cómo estos problemas, y las ideas matemáticas que movilizaron y que fueron tan fértiles en la época, podrían ser tomadas en cuenta para el trabajo en aulas del nivel superior.
En el segundo capítulo nos apoyaremos en una indagación en bibliografía especializada en probabilidad (especialmente en aquella que se utiliza en nuestra casa de estudios) con la intención de poner en consideración aspectos teóricos en torno a algunos asuntos que esta tesis se propone abordar. En particular, hacemos una sumarización de las técnicas empleadas en las diversas demostraciones de la Ley de los Grandes Números y del Teorema Central del Límite.
Estos elementos se pondrán en juego en el tercer capítulo, donde desarrollaremos un marco teórico general que enmarca nuestro posicionamiento didáctico. En este capítulo consideraremos algunos aportes específicos de la didaçtica de la probabilidad y la estadística, y realizamos un análisis didáctico-matemático de una propuesta didáctica, con la intención de que pueda ser de utilidad para docentes, tanto de nuestra facultad como de otras.
Finalmente, el capítulo de conclusiones trataremos de atrapar algunas cuestiones globales que habremos desarrollado, y dejar expresados lineamientos para profundizar o ampliar lo estudiado.
Para concluir, queremos enfatizar particularmente que nuestra experiencia de trabajo en la materia Probabilidad y Estadística (M)/Probabilidades (D), Probabilidad y Estadística (C), Estadística para la Licenciatura en Química, entre otras, es un punto de apoyo para la reflexión sobre lo propuesto en esta

tesis de licenciatura. Es conveniente señalar, a estos fines, que algunas ideas, conclusiones o propuestas tomaron en consideración esta experiencia, y nos permite sostener como premisa que la actividad docente es, sin lugar a dudas, también una actividad de producción de conocimiento.

Capítulo 2

Los problemas

> En definitiva, la probabilidad es sentido común reducido a cálculo
>
> *Laplace, 1810*

Una mirada que atraviesa el desarrollo de esta Tesis es cómo los asuntos matemáticos que pretenden ser abordados desde la propuesta son productos ya elaborados en respuesta a determinados problemas intramatemática y extramatemáticamente complejos. Fueron necesarios procesos complejos y temporalmente extensos, dados hacia adentro de la comunidad de producción matemática (de los cuales las discusiones entre matemátiques fueron particulares protagonistas) no solo en términos la pertinencia de las soluciones a dichos problemas, sino de la producción y desarrollo de conocimiento relevante para entenderlas, mejorarlas y profundizarlas. Se parte de la pregunta de cómo tomar en cuenta estos elementos (además de los que proveen los libros de texto), no desde el punto de vista de un relato, sino desde la propia sustancia de los conceptos en cuestión, de modo que considerarlos en un diseño didáctico que pueda ofrecer a les estudiantes nuevas oportunidades de construir conocimiento.

El abordaje de una investigación que incorpore elementos para pensar una propuesta didáctica en torno a dos grandes teoremas como lo son la Ley de los Grandes Números y el Teorema Central del Límite nos lleva a pensar en cómo fueron evolucionando a lo largo del tiempo hacia adentro de la comunidad científica y no científica. Además de buscar una suerte de cronología que nos permita dar cuenta de algunos hitos en este desarrollo, nos interesa indagar qué tipo de problemas han venido a resolver estos resultados, pero más: intentar encontrar, en la génesis histórica de estos conceptos, aspectos de la complejidad de su desarrollo que puedan ser de interés para esta tesis.
A esos fines, en este capítulo nos proponemos realizar un recorte sobre cómo los problemas

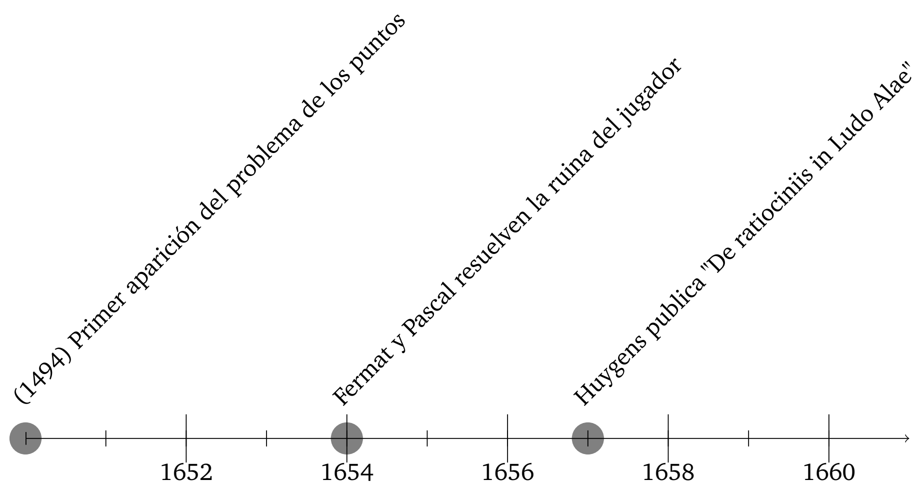
a) *determinar* la probabilidad de ocurrencia de un fenómeno aleatorio para el que no se dispone de un modelo probabilístico,
   y

b) *determinar* la distribución del error aleatorio de un conjunto de mediciones obtenidas a partir de un instrumento calibrado

fueron abordados entre 1700 y 1840. Si bien el termino *determinar* puede resultar un tanto pretencioso para la mirada actual de que propone la estadística, optamos por plantear los problemas siguiendo su formulación original. Veremos también cómo el tratamiento de estos problemas posibilitó el desarrollo de teoría y técnica estocástica que modernamente es relevante, tanto para la matemática en sí como para la educación matemática. Asimismo, a partir de este recorte, queremos extraer elementos interesantes para el desarrollo de este trabajo. Al principio de cada sección, incorporaremos una línea de tiempo que puede ser útil para tener de referencia a lo largo de la lectura.
En este capítulo tomaremos como referencias principales los libros [34] y [41].

2.1. El problema del azar

Algunas referencias a los inicios de la probabilidad



Aunque los juegos de azar[1] parecen haber comenzado a practicarse hace al menos 4000 años[2] y haber sido populares en Grecia y Roma, no se dio, como por ejemplo en el caso de la geometría, un interés o estudio con relación a ellos en particular, y al azar en general, sino hasta principios del 1500, al respecto de los varios problemas que modernamente se conocen como .El problema de los puntos":

> Dos jugadores juegan un juego justo, y deciden que el premio será para aquél que gane seis rondas. Supongamos que el juego debe detenerse antes de que alguno gane. ¿Cómo debería dividirse el premio? [53]

Grandes matemátiques como Fray Luca de Pacioli en 1494 o Nicolo Tartaglia en 1556 propusieron soluciones, sin contar con un conjunto de conocimientos específicos que permitieran estudiarlo.

[1]Del árabe “al-zar”, flor, por la que se pintaba en una de las caras del dado

[2]El primer juego del que se tiene conocimiento, y que aún hoy es practicado en España, es el de tirar unos pequeños huesos llamados ástragalos, que hacían sus veces de dados

Quienes finalmente pudieron dar una respuesta satisfactoria fueron Pierre de Fermat y Blaise Pascal, considerados fundadores del estudio de la probabilidad a mediados del 1600. La solución impactó en la comunidad de aquel momento, ya que consistía en conjeturar sobre una base combinatoria al respecto de eventos que ya no ocurrirían (puesto que, como indica el enunciado, el partido es interrumpido). Según Batanero y Godino, existían dificultades para concebir la idea de suceso aleatorio. Al respecto, señalan que:

> “una de las posibles causas de estas dificultades (...) era el uso que se hacía de [los dados] en ceremonias religiosas para que la divinidad mostrara su voluntad. Cualquier intento de prever el resultado del lanzamiento podía ser interpretado como una pretensión de adivinar la acción de la deidad correspondiente, y este acto de impiedad podría tener mala suerte (...) Kendall (1978) incorpora como otros posibles elementos a tomar en consideración: la ausencia del desarrollo de un álgebra combinatoria y la existencia de barreras morales o religiosas para el desarrollo de la idea de aleatoriedad y azar (P. 30 y 31) [6]

Estos elementos nos hacen poner en consideración la existencia de una distancia, desde el punto de vista epistemológico, entre la concepción de que sobre determinados sucesos no se sabe cuál de ellos ocurrirá, y la cuantificación de esa incerteza, incluso cuando los sucesos en cuestión son parte de juegos cuyas reglas se encuentran perfectamente definidas. Luego de que Fermat y Pascal compartieran su solución, «todo París estaba hablando del descubrimiento de una nueva ciencia» [53] En 1657, el matemático Cristian Huygens publicó "De ratiociniis in Ludo Alae.en 1657, donde no solo exhibe las soluciones a diversos problemas originados en juegos de azar, entre ellos la Ruina del Jugador , sino que

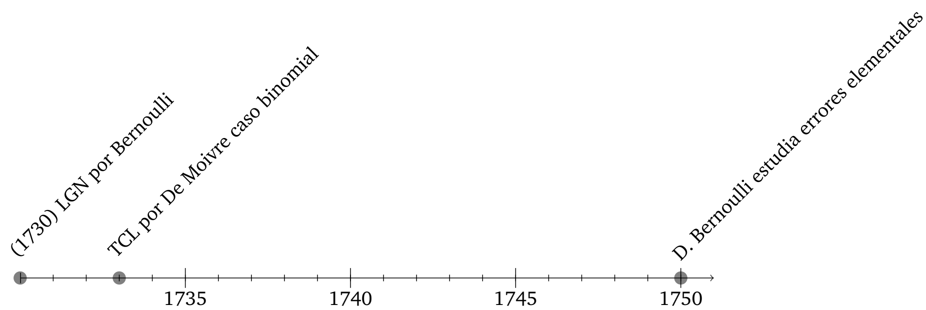
> El pequeño tratado se convirtió en el primer trabajo impreso sobre cálculo de probabilidades y referencia básica para los autores que en los inicios del siglo XVIII irrumpieron tan arrolladoramente en la consolidación del mismo. [52]

Ahora bien: la probabilidad y el azar se encontraban circunscriptos al ámbito de los juegos de azar. En términos actuales, trabajaban en espacios equiprobables, haciendo desarrollos combinatorios para poder contar. Llevarla a otros terrenos constituyó un salto difícil y considerable, de la mano de Jakob Bernoulli y de la primera formulación de la Ley de los Grandes Números. Como lo refiere Anders Hald:

> En los tiempos de Bernoulli, la Doctrina de posibilidades, nombre que recibía la probabilidad, tenía como objetivo el cálculo de la esperanza de ganancia de un juego de apuestas. Gracias a las simetrías de estos juegos, todos los sucesos posibles eran considerados igualmente probables y permitieron la definición clásica de probabilidad como el cociente entre el número de casos favorables al total de casos posibles. La gran visión de Bernoulli fue extender esta doctrina de posibilidades a una teoría de probabilidades que permitiera abordar eventos inciertos (...) en los que la definición clásica es imposible de usar, ya que los eventos en cuestión dependen de numerosas causas desconocidas” [41]

El trabajo de Bernoulli

> Todas las cuestiones del cálculo de probabilidades pueden reducirse a una hipótesis simple: que una cierta cantidad de bolillas de colores diferentes han sido mezcladas en una urna, de la cual se extrae una cierta cantidad al azar en un cierto orden y en una cierta proporcion
>
> *Condorcet, en el prefacio de "Mémoire sur les probabilités" de Laplace, en 1780, citado por Daston en [26]*



Parafraseando a Daston en [26], la "Doctrina de posibilidades" tenía como objetivo ofrecer un discurso racional que funcionara para sostener una cierta manera de hacer ciertas cosas (aceptar una hipótesis científica, realizar una determinada inversión, etc), pero no contaba con un discurso propio específico constituido hacia dentro de la comunidad matemática, De hecho, es en interrelación con estos problemas a resolver, y con desarrollos matemáticos, que la teoría de probabilidades se ha desarrollado a lo largo de la historia.

> "La teoría clásica de probabilidades era una disciplina de la matemática en un sentido amplio, construida sobre un cierto consenso alrededor de qué era lo "razonable", y que a su vez debía funcionar como una asistencia al "sentido común" a partir de incorporar cálculos durante la toma de decisiones". [26]

Entre la gran cantidad de matemátiques con apellido Bernoulli que se dedicaron, entre otras áreas, al desarrollo de la matemática, cinco de elles hicieron importantes contribuciones a la probabilidad y estadística. El primero, Jakob, escribió su famosa Ars Conjectandi, traducido como El arte de conjeturar, libro de gran importancia en la historia de la probabilidad. Es interesante observar el sentido que el término "conjeturar.ªdquiere dentro del libro: "designamos como el arte de la conjetura o de la suposición al procedimiento de medición más exacta de las probabilidades de las cosas". Este libro, en particular, es de los primeros en ocuparse de la tarea de definir la probabilidad:

> “La probabilidad es un grado de certeza, y se diferencia de ella como la parte del todo (...) si la certeza total o absoluta, que designamos por medio de (...) 1, consta de cinco partes, de las cuales tres son favorables al acaecimiento presente o futuro de cierto suceso (...) en tal caso diremos que el suceso tiene $\frac{3}{5}a$ o $\frac{3}{5}$ de certeza” [7].

Como podemos leer, la noción de probabilidad de Bernoulli está asociada a la idea de certeza, y se calcula como actualmente se calcula la probabilidad en espacios equiprobables.
Por otra parte, el autor plantea un problema:

> “Pero aquí parece que, para la formación correcta de suposiciones sobre una cosa cualquiera, tan solo es preciso averiguar exactamente el número de estos casos [aquellos en los que esta cosa ocurre, y aquellos totales] y luego determinar en qué medida unos casos pueden ocurrir más fácilmente unos que otros. Pero aquí parece residir precisamente la dificultad, pues esto solo es posible en un número mínimo de fenómenos y casi exclusivamente en los juegos de azar, pues (...) fueron diseñados por sus creadores de tal manera que los números de casos en los que el resultado es ganancia o pérdida están determinados previamente y son conocidos, y todos los casos tienen la misma facilidad de ocurrencia. Pero este no es el caso para los muchísimos otros fenómenos que depende de las fuerzas de la Naturaleza o del arbitrio de los hombres” [7].

En términos actuales, es el primer problema propuesto al principio del capítulo: determinar la probabilidad de ocurrencia de un fenómeno para el cuál no se dipone de un modelo probabilístico. Cabe mencionar que este problema suele plantearse en cursos de estadística como el caso más sencillo de inferencia, luego del desarrollo de una gran cantidad de teoría probabilística.
Jakob Bernoulli, para la solución, propone:

> “Pero hay abierto otro camino para hallar lo que buscamos y averiguar, por lo menos a posteriori, e.d., por medio del resultado de lo que se ha observado en numerosos casos en ejemplos similares, lo que a priori no podemos determinar. Para ello hay que admitir que todo suceso individual puede acontecer o no acontecer en el mismo número de casos, al igual que se observó antes, en un estado igual de las cosas (...) esta forma empírica de determinación por medio de observaciones no es nueva ni inusual (...) le resulta evidente a toda persona que, para enjuiciar un suceso cualquiera, no basta llevar a cabo una observación u otra, sino que es necesario un número grande de observaciones (...) pero si bien esto, por la naturaleza de la cosa, lo entiende cualquiera, la prueba fundada en principios científicos no es evidente, por lo que estoy obligado a ofrecerla en este lugar” (P. 401). [7]

Así planteado, el problema de Jakob Bernoulli queda en una primera zona: demostrar que es deseable realizar un gran número de observaciones para inferir información sobre la probabilidad de ocurrencia de un determinado fenómeno. No obstante, él amplía:

> “Pero creería que mi contribución es muy pequeña si me quedara en la prueba de este único hecho que todo el mundo conoce. Más bien haya que tomar en consideración algo en lo que

nadie haya pensado todavía. Se trata de investigar si, por medio del incremento de observaciones, crece constantemente la probabilidad de que el número de observaciones favorables alcance, respecto del número de observaciones desfavorables, la proporción verdadera, y si esta probabilidad supera en definitiva todo grado de certeza.” [7]

Es decir: la ley débil de los grandes números para variableas aleatorias con distribución Bernoulli de parámetro p, donde p denota la proporción en cuestión.. El detalle de la demostración del matemático puede encontrarse en el Capítulo VI del libro de Uspensky [58]. Al respecto de ella, nos interesa resaltar dos asuntos que son de interés para este trabajo:

1. Bernoulli, en su definición de probabilidad como certeza, asume que los espacios muestrales son equiprobables y por lo tanto.

2. La demostración utiliza solamente, aunque de manera ingeniosa, la fórmula de lo que hoy conocemos como distribución binomial.

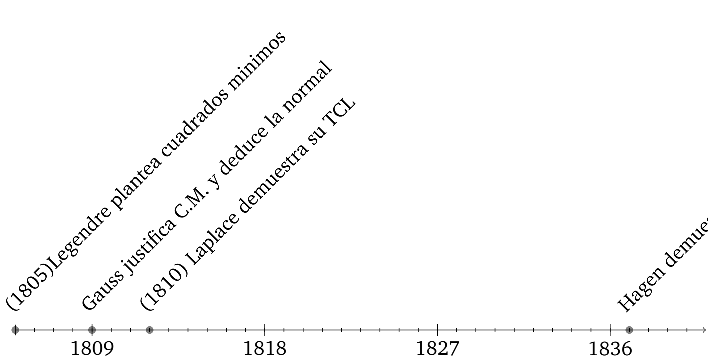
Como mencionaremos en el siguiente capítulo, estas consideraciones sobre las hipótesis que el matemático utiliza para realizar la demostración nos permiten considerar posibles extensiones para la validez de su razonamiento, es decir: ¿qué axiomas mínimos serían necesarios para, sin necesidad de trabajar en un espacio muestral equiprobable, arribar a la fórmula de la binomial y, con ello, poder extender la validez de esta demostración? Amen de esta discusión, nos parece interesante cómo este resultado, aún en el caso restrictivo de un espacio muestral equiprobable, permite afirmar la convergencia de las frecuencias relativas sin verla como un resultado particular de Ley Débil.
Observemos que en la formulación misma del problema, y la de la solución, la idea de determinar la probabilidad experimentalmente tiene un rol central, tanto es así que Bernoulli no quiso publicar su trabajo -y de hecho fue una publicación póstuma-, aparentemente a causa de que la cantidad de repeticiones necesarias (para estimar con un error pequeño) a la que él llegó con su demostración era tan grande que resultaba inaplicable, por lo que lo consideró irrelevante para sus contemporáneos. Afortunadamente, se equivocó: tal fue el revuelo de su trabajo, que De Moivre dedicó un buen tramo de su famosa "Doctrine of chances” en mejorar el trabajo de su predecesor. Lo que no sabría es que, en este esfuerzo, daría una primera aproximación al Teorema Central del Límite, demostrando que la curva que modernamente conocemos como "Gaussiana.ªproxima bien a la funcion de probabilidad puntual de la binomial, lo cual actualmente es visto como una aplicación del Teorema Central del límite, pero que en los hechos es la primera aparición de la Gaussiana en la literatura. Un detalle de este trabajo puede leerse en el [41].
Antes de seguir avanzando, conviene observar, por una parte, que la resolución planteada por Bernoulli es modernamente válida, y que no requirió, por su parte, del desarrollo de una teoría moderna de probabilidades para ser comprendida en su momento.
Por último, el hecho de que Bernoulli descartara su trabajo por su inaplicabilidad es interesante, puesto que destaca la relación que existía para el matemático entre el resultado probabilístico y la posibilidad de realizar una experimentación cuyos resultados pudieran ser analizados.

Quisiéramos, entonces, avanzar sobre el desarrollo del Teorema Central del Límite: hasta aquí, la aproximación normal de la binomial es casi un resultado técnico, significativo en la medida de que subsana el problema de la inaplicabilidad encontrado por Bernoulli. En la sección siguiente nos proponemos abordar los siguientes pasos del Teorema Central del Límite, como refiere Hans Fischer:

> En sentido estricto, en realidad no hay que referirse [antes de 1810] al teorema central del límite en relación con sumas de variables aleatorias independientes, sino a un teorema central del límite caso por caso. No obstante, después de 1810, cuando Laplace utilizó distribuciones normales para presentar aproximaciones que eran válidas en situaciones decididamente generales, la afirmación relativa a la existencia universal de una "ley de frecuencia"similar a e^{-x^2} para sumas de grandes números de variables aleatorias independientes adquirió el estatus de ley natural a los ojos de casi todos los probabilistas del siglo XIX [34]

2.2. El problema de la medición

> No he tenido ninguna suerte con la distribución gaussiana
>
> *Lévy, 1970*



Cronológicamente hablando, nuestro primer hito se encuentra en el trabajo de Laplace, quien se dedicó a estudiar, entre 1780 y 1810, si las posiciones de los planetas y de determinados cometas estaban influenciadas por algún fenómeno celeste desconocido. Uno de sus grandes objetos de estudio fue el ángulo de inclinación de 97 cometas respecto de la eclíptica[3]. Habiendo medido el ángulo de cada uno, y obteniendo un promedio observado de aproximadamente 51,98 grados, Laplace se preguntó por la probabilidad de que el promedio *cayera* en el intervalo $[50-1,98; 50+1,98]$. Solo una vez que Laplace hubo producido, luego de más de 30 años de trabajo, su versión del Teorema Central del Límite, pudo determinar que dicha probabilidad era aproximadamente $\frac{1}{2}$, y concluyó que los cometas no se encontraban influenciados por ningún fenómeno celeste en particular. En el libro de Hans Fischer ([34]), y en

[3]Curva que describe la trayectoria del sol, tomando un sistema de referencia centrado en el planeta tierra

el de Hald ([41]), se puede encontrar una referencia precisa a cómo Laplace fue atacando el problema con diferentes técnicas. Sí nos interesa destacar que el matemático, asumiendo que los ángulos seguían una distribución uniforme entre 0 y 100 grados, primero intentó producir una fórmula para calcular de manera exacta esta probabilidad. Ante la imposibilidad de aplicar la formula obtenida, migró a diferentes estrategias de aproximación. De Moivre ya había estudiado, para el caso de estimar la probabilidad de la suma de los resultados de tirar n dados de k caras, la probabilidad de que la suma tome valores en una cierta región, utilizando funciones generadoras de probabilidad[4]. Lo que fue clave para que Laplace lograra su versión del TCL fue la modificación de las funciones de generadoras de probabilidad, sustituyendo el término t por e^{ix}, constituyendo una referencia crucial de los trabajos posteriores de Liapunov, que en 1908 dio la primer demostración moderna del TCL utilizando, justamente, funciones características. Un detalle más explícito puede encontrarse en A History of CLT. El autor, además, refiere que

> “Laplace en ningún momento estableció un teorema general que se corresponda con el Teorema Central del Limite en el sentido actual. Él solo trabajó con problemas particulares relativos a la aproximación de sumas o combinaciones lineales de un gran número de variables aleatorias (en muchos casos, errores de observación)”. [34]

Esta referencia nos permite interiorizarnos en un problema que entrelaza a Laplace, Legendre y Gauss, y en torno al cual se consolida la difusión del Teorema Central del Límite, y de la distribución normal. Nos interesa señalar que, tanto en el trabajo de De Moivre como en el de Laplace, la aparición de la famosa e^{-x^2} responde a la supresión de términos de orden superior a 2 empleados para la aproximación de la densidad de la suma de N variables aleatorias, mientras que en el trabajo de Gauss, el surgimiento de la normal se debe a un razonamiento que encontramos muy interesante, y que desde el punto de vista estocástico tiene su interés también. En términos actuales, Gauss propone un modelo de posición y muestra que si la función de verosimilitud se maximiza en el promedio, entonces la distribuciń central es la densidad normal centrada. La demostración de este resultado se incluye en la Sección 2.3.1

Ya desde el año 1750, mucho antes de que Laplace hubiera demostrado su versión del TCL, era objeto de estudio el problema general de hallar un vector ψ que ajuste un modelo experimental de la forma

$$d + \varepsilon = A\psi, \tag{2.1}$$

donde A es una matriz dada, $d = (d_1, \ldots, d_n)$ es un vector de n observaciones, y $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ es un sumando aleatorio n-dimensional, denominado en este trabajo error observacional. donde A Si bien modernamente los modelos de posición se expresan con el sumando aleatorio en el miembro derecho, en [34] se recupera el uso epocal de dicho sumando. El hecho de que ε se asuma simétrico hace que una y otra formulación sean equivalentes.

---

[4]Dada una sucesión $\{a_k\}_{k\in\mathbb{N}}$, su función generadora asociada es el polinomio $P(t) = \sum_{k=1}^{\infty} a_k t^k$. Observar que si escribimos, para lo que hoy conocemos como "Función generadora de Momentos”, su desarrollo en serie de Taylor en torno a $x = 0$, obtenemos una expresión análoga a la de función generadora de probabilidad. Es conveniente observar que, en estas épocas, las series de potencias eran de extrema utilidad para numerosos ámbitos de la matemática

Para fijar ideas -y el caso base con el que trabajó Gauss-, observemos que un caso particular de este problema general es considerar, para cada i, $d_i + \varepsilon_i = \psi$, teniendo ε_i la misma distribución, para todo i. Este caso recibe el nombre de *el caso de la observación directa*, en el que el valor real (ψ) difiere del i-ésimo valor observado en un término aleatorio aditivo denominado *error*. Diversos métodos para la estimación de ψ fueron desarrollados a partir de 1750 (Laplace, de hecho, propuso dos diferentes, uno en 1788, conocido como "método de promedios", y otro en 1799, conocido como "método de mínimas desviaciones absolutas"), hasta que, como refiere Hald ([41]), en 1805 Legendre propuso el método de cuadrados mínimos[5].
Legendre dio argumentos intuitivos para sostener que el método de cuadrados mínimos, originalmente considerado para el modelo (2.1), era muy bueno para estimar ψ. Notemos que para el caso de observación directa, este método estima a ψ con el promedio de las observariones d_i. Tanto Laplace como Gauss se propusieron demostrar *la optimalidad* [6] del método de cuadrados mínimos. Laplace utilizó, para el caso general presentado en (2.1), su recién demostrado Teorema Central del Límite. En cambio Gauss, para el caso de la observación directa, partió de la siguiente hipótesis:

> "Se ha considerado ciertamente como un axioma la hipótesis de que si alguna cantidad ha sido determinada por varias observaciones directas, realizadas en las mismas circunstancias y con igual cuidado, la media aritmética de los valores observados proporciona el valor más probable [a los valores observados] (Gauss, 1809 [37], p. 258, citado por Hald [41]) [37].

Es decir, postula que $\widehat{\psi}$ definido por $\widehat{\psi} = \frac{\sum_{i=1}^{n} d_i}{n}$ es el valor que maximiza lo que hoy conocemos como la función de verosimilitud en un modelo de posición. Este hecho permite que pueda deducir que si ϕ denota la densidad de ε y $\widehat{\psi}$ es el valor de x que maximiza

$$\phi(x - d_1) \dots \phi(x - d_n), \tag{2.2}$$

entonces ϕ es lo que hoy conocemos como densidad Gaussiana. El detalle de esta demostración se encuentra al final del capítulo.

Luego, para el caso general planteado en (2.1), asume que ε_i se distribuye normalmente, pudiendo definir una función de verosimilitud en el sentido moderno. Invocando nuevamente un argumento de máxima verosimilitud concluye que el método de cuadrados mínimos coincide con los valores que maximizan la verosimilitud.

Fue objeto de crítica en la época la asunción de que el promedio es el mejor estimador de la magnitud desconocida, la propuesta de que ε_i se distribuye normalmente y también la factorización presentada en (2.2). No obstante, la claridad y concisión de estas ideas tuvo una aceptación muy amplia, tanto que

[5]Omitimos una discusión que da Hald en torno a que, si bien la primera publicación donde aparece cuadrados mínimos es la que mencionamos, "Laplace escribió que Legendre fue el primero en publicar el método, pero que debemos a Gauss la justicia de observar que él había tenido la misma idea varios años antes, que la había utilizado y que la había comunicado a varios astrónomos."Hald, [41] P.50

[6]A lo largo de estas lineas haremos un uso vago de la palabra optimalidad, en consonancia con el tratamiento que recibe en los libros de referencia de este capítulo ([41], [34].

aunque Gauss produjo poco tiempo después otra justificación de la optimalidad que no implicaba deducir la distribución de ε, su primer solución se difundió ampliamente. Un curioso ejemplo en la historia de la matemática en el que una primer, pese a sus falencias matemáticas, pregna en el discurso de la época dentro de la comunidad científica de manera significativa aún habiendo presentado poco tiempo despés una solución superadora. Uno de los atractivos que también recibió esta solución es que permitía ubicar el problema de conocer una determinada *curva de errores*, es decir, el gráfico de la función de densidad del error, en términos de un problema paramétrico, al incorporar lo que hoy conocemos como la varianza del error. De esta forma, llegamos al conocido modelo de posición y escala para las mediciones, que asume que $d \sim N(\psi, \sigma^2)$, uno de los pilares de la inferencia estadística.

Persistía, junto al eco de las críticas, la pregunta de si esta curva del error propuesta resultaba consistente con la manera en la que se distribuían las observaciones (en consonancia con el status de la teoría de probabilidades en aquella época) y si podían subsanarse las limitaciones de la demostración de Gauss.

Hagen y Bessel se abocaron a esta tarea. Como refiere Hald,

> "En 1818, en el estudio de las ascensiones y declinaciones de estrellas, Bessel hizo una comparación entre las frecuencias relativas de los residuos de observaciones directas que caían en un determinado intervalo, y las respectivas probabilidades calculadas según la ley de errores de Gauss, encontrando una buena correspondencia entre ambos conjuntos de valores. [41]

Por otro lado, Hagen (discipulo de Bessel) postula la denominada hipotesis de los errores elementales, según la cual

> "El error resultante de cualquier medición es la suma aritmética de una cantidad infinitamente grande de errores elementales ["elementäre Fehler"], todos del mismo módulo, donde cada uno puede ser positivo o negativo" [34]

En [34] se encuentra la demostración de Hagen de la Ley Gaussiana, que en consonancia con la hipótesis de errores elementales, considera la suma de $2n$ variables aleatorias X_i a valores $-\frac{\Delta x}{2}$ y $\frac{\Delta x}{2}$, cada uno con probabilidad $\frac{1}{2}$.

Conviene señalar que, en 1838 Bessel generalizó la hipótesis de errores elementales, detalle que puede leerse en [34].

Antes de terminar este breve recorrido, nos interesa observar que la idea de estudiar la distribución del error observacional no fue original de Legendre. De hecho,

> "Daniel Bernoulli [1778] ya había insinuado la idea de que cualquier error observacional es producto de la suma de una gran cantidad de errores muy pequeños (...) y en 1780 discutió profundamente las "aberraciones" de relojes de péndulo a través de un modelo binomial, en el que asumía una acumulación de desviaciones equiprobables e iguales en modulo, aunque no necesariamente en signo" [34]

No obstante, se considera que Hagen desconocía el trabajo realizado por Daniel Bernoulli. De hecho, este trabajo casi no tuvo repercusión en su época.

2.3. Consideraciones parciales

En el presente capítulo hemos abordado, realizando un recorte, las versiones de los problemas arriba referidos en según diferentes autores entre el 1700 y 1840, y sus respectivas soluciones, ya sean parciales o totales. Para tener una perspectiva más completa de esto, sería necesario realizar un estudio del uso de la probabilidad inversa y el rol de los argumentos bayesianos a lo largo de este periodo, puesto que como muestra el libro de Hald, estos asuntos fueron parte, no solo de los conocimientos en construcción en aquella época, sino del desarrollo de de argumentos, intuiciones y problemas relevantes para desarrollo de los conceptos que abordamos en esta tesis, y de otros más en general que son interesantes tanto para la probabilidad como para la estadística. En particular, y como mencionamos más abajo, la deducción de la normal por Gauss se apoya en un razonamiento que apela a la relación entre probabilidades a priori y a posteriori. Por su parte, no solo Laplace en 1774 ya había hecho considerables estudios en torno a la relación entre estas probabilidades, sino que había formulado lo que hoy conocemos como "Teorema de Bayes” y abordado discusiones relevantes en torno a la Ley de los Grandes Números de Bernoulli tomando en consideración argumentos apoyados en probabilidad inversa. Finalmente y con relación a este asunto, nos interesa mencionar que razonamiento someramente esbozado en torno al problema de los cometas que mencionamos, permite ubicar el trabajo de Laplace con estrecha relación a la noción de intervalo de confianza.

También sería relevante incorporar el trabajo de Poisson, el cuál propuso una generalización a la Ley de los Grandes Números de Bernoulli, relajando la hipótesis de equidistribución e independencia, y discutiendo ejemplos de no aplicación de la Ley (destacando, entre ellos, el de la variable aleatoria que actualmente recibe el nombre de Cauchy).
Tampoco hemos incorporado las discusiones posteriores al trabajo de Hagen, las cuales son particularmente interesantes e involucran a matemátiques y estadistas como Quetelet y Galton, este último particularmente famoso por el diseño del tablero de Galton, o Quincunx, diseñado para simular distribuciones binomiales, y su relación con el Teorema Central del Límite. De hecho, nos inspiramos en el trabajo de este último para parte de la elaboración de la secuencia didáctica. Dar cuenta de manera precisa del entramado de conocimientos que hemos esbozado en estas últimas líneas supone un despliegue de más largo aliento que, si bien excede el trabajo de esta tesis, pueden ser una vía para profundizarlo.

A partir de lo desplegado en este capítulo, aparecen algunos asuntos que nos interesa señalar.

- El desarrollo de conocimiento, digamos probabilístico, guarda una relación muy estrecha con la solución de determinados problemas que hoy en día ubicamos en el terreno de la estadística, específicamente el de la inferencia estadística (ya sea puntual o por intervalos). Por otra parte, la pertinencia de estas soluciones estuvo históricamente determinada por la puesta a prueba de las mismas, además de por criterios intrínsecamente matemáticos. Por tanto, puede ser interesante considerar un enfoque que se proponga no necesariamente reconstruir, pero sí articular, a la hora de abordar determinados saberes en el aula, sentidos pertenencientes a ambas disciplinas a propósito de problemas que las requieran genuinamente en su abordaje.

- Más en particular, el tratamiento histórico de los dos problemas mencionados al principio del capítulo estuvo en estrecha relación con el desarrollo de varios de los conceptos matemáticos que se abordan en la secuencia de que es objeto esta Tesis. Esta observación nos permite ponderar la posibilidad de diseñar nuevos problemas para el aula que permitan, a partir de su propuesta, la generación de condiciones para que les estudiantes se aproximen a algunos conceptos probabilísticos y/o estadísticos, en particular de la Ley Débil de los Grandes Números y del Teorema Central del Límite.

- Algunas de las soluciones dadas a estos problemas -pero que no hemos desplegado aquí- involucran la articulación de conocimientos y técnicas pertenecientes a áreas diversas (teoría de números, análisis complejo, cálculo diferencial e integral) las cuales no necesariamente son pertinentes de abordar en un curso introductorio de probabilidad y estadística. En el caso de la prueba de Jakob Bernoulli, por su parte, si bien no apela a conocimientos que se ubiquen en otras áreas, el tipo de trabajo realizado puede resultar cognitivamente demandante si no se tiene particularmente en claro el esquema de la demostración y el sentido de los diferentes lemas que son necesarios para llevarla adelante, redundando en un complejo esfuerzo cognitivo que bien puede ser puesto en juego en otro tipo de actividades o desarrollos.

A partir de estas consideranciones, en adelante tendremos como norte abordar la pregunta de:

> ¿Cómo tomar en consideración los elementos desarrollados en este capítulo, en torno a los problemas de "determinar la probabilidad de ocurrencia de un fenómeno para el que no se conocen sus leyes probabísticas” y "determinar el comportamiento del error aleatorio de un conjunto de mediciones” para elaborar una propuesta didáctica que contenga la intención de promover espacios de trabajo matemático en aulas de matemática donde se construyan conocimientos estocásticos en general, y en torno a la Ley de los Grandes Números y al Teorema Central del Límite en particular?

En función de lo desarrollado en este capítulo, sostenemos que si bien estos problemas son matemáticamente relevantes y movilizaron conocimientos relativos a lo que interesa estudiar, su fertilidad y potencialidad dentro de un aula de matemática dependen de la posibilidad de adaptarlos en el marco de una propuesta didáctica que permita sostener su tratamiento. En este sentido, nos distanciamos de una posición que pretenda un intento de homologación o recreación histórica con la intención de abordar estos conceptos. No obstante, es necesario señalar cómo, a partir de este estudio, se obtuvieron referencias concretas no solo relativas a la interrelación histórica entre la solución de estos problemas y el desarrollo de determinados conceptos, sino a la constitución de un punto de observación didáctico-matemático para posteriores desarrollos y tratamientos de estos asuntos. En este sentido, para producir la adaptación en cuestión necesitamos avanzar en el estudio de producciones matemáticas y didáctico-matemáticas en el área de probabilidad y estadística, de lo que nos ocuparemos en los siguientes capítulos.

2.3.1. Deducciones de la curva normal

Vamos ahora a presentar la deducción de la curva normal hecha por Gauss, [34], en lenguaje moderno. Para ello, denotemos con ϕ a la curva del error observacional (en términos actuales, la densidad de la distribución del error). Utilizaremos M_i en lugar de d_i para denotar las mediciones. La hipótesis de Gauss es que $L(x) = \phi(x - M_1)...\phi(x - M_n)$ se maximiza en $x = \frac{\sum_{i=1}^{N} M_i}{n}$. Modernamente, $L(x)$ es lo que conocemos como función de verosimilitud. Asumiremos, como lo hizo Gauss, que es una función simétrica respecto al origen.

Tomando logaritmo, el matemático obtiene la expresión[7]:

$$\ln(L(x)) = \ln(\phi(X - M_1)) + ... + \ln(\phi(X - M_n)) \tag{2.3}$$

y luego, derivando respecto de x y utilizando que $x = \frac{\sum_{i=1}^{N} M_i}{n}$ maximiza a $L(x)$ por hipótesis, resulta que dicho x es solucion de

$$\frac{\phi'(x - M_1)}{\phi(x - M_1)} + ... + \frac{\phi'(x - M_n)}{\phi(x - M_n)} = 0. \tag{2.4}$$

Ahora bien: como la asunción de Gauss es para cualquier conjunto de observaciones, y no depende del tamaño de n, Gauss en particular plantea considerar el caso particular donde $M_i = M_1 - N.n$, para $i = 2, \dots n$, donde N es un valor cualquiera M_1 es arbitrario. El promedio correspondiente a estos valores esta dado por $M_1 - (n-1)N$ y reemplazando en (2.4), llegamos que a ϕ debe satisfacer la siguiente condición:

$$\frac{\phi'((n-1)N)}{\phi((n-1)N)} = (1-n)\frac{\phi'(N)}{\phi(N)}.$$

Ahora bien, esto es equivalente a

$$\frac{\phi'((n-1)N)}{\phi((n-1)N)(n-1)} = -\frac{\phi'(N)}{\phi(N)}.$$

Luego, si multiplicamos por $\frac{1}{N}$, tenemos que

$$\frac{\phi'((n-1)N)}{\phi((n-1)N)N(n-1)} = -\frac{\phi'(N)}{\phi(N)N}.$$

[7] Le Cam, en 1986, explica que el planteo de Gauss tuvo notables detractores que, por su parte, señalaron diversas fallas en el argumento. Una de las críticas más fue que la densidad de las observaciones no debería ser del tipo $f(x-\theta)$ sino más generalmente $f(x,\theta)$. Bertrand mostró que si el promedio de las observaciones es el estimador de máxima verosimilitud, entonces $f(x,0)$ no tiene por que ser Gaussiano, aunque si de la familia exponencial. También critica el principio del promedio aritmético, explicando que las observaciones anómalas deben tener menos peso que las otras.

Sustituyendo $\Delta = N(n-1)$, se tiene que para todos $n \in \mathbb{N}, N \in \mathbb{R}$

$$\frac{\phi'(\Delta)}{\phi(\Delta)\Delta} = -\frac{\phi'(\frac{\Delta}{n-1})}{\phi(\frac{\Delta}{n-1})\frac{\Delta}{n-1}}$$

Luego, fijado $z \in \mathbb{R}$, tomando $N = \frac{z}{n-1}$, resulta

$$\frac{\phi'(z)}{\phi(z)z} = -\frac{\phi'(\frac{z}{n-1})}{\phi(\frac{z}{n-1})\frac{z}{n-1}}$$

De lo que si

$$\lim_{u\to 0} \frac{\phi'(u)}{\phi(u)u} = K,$$

concluimos que, para todo $z \in \mathbb{R}$,

$$\frac{\phi'(z)}{\phi(z)z} = K.$$

La solución de esta última ecuación diferencia está dada por la densidad de una normal de esperanza 0 y varianza K.

Capítulo 3

La Matemática y el Aula

Un aula -vista hacia dentro de una institución, con sus docentes y sus estudiantes- es un territorio donde circulan conocimientos de variada índole. Muchos de estos conocimientos están referidos a ciertos *saberes* que existen en determinadas comunidades científicas, las cuales los expresan mediante producciones de naturaleza diversa. Cuando se trata de poner a circular algunos de estos saberes en instancias de formación de estudiantes, unas producciones particulares cobran relevancia: los libros de texto.

Tanto en la etapa de conformación de una determinada asignatura, como en el proceso de dictado de la misma, estos libros son una referencia: allí se encuentran expresados **consensos** relativamente amplios en torno a los asuntos matemáticos que se tratan en ellos (definiciones, nombres, objetos) y se encuentran por doquier **decisiones** de les autores en torno a la estructuración, organización y enfoque de sus contenidos.

En este capítulo, hemos realizado una incursión en parte de la bibliografía especializada en probabilidad y estadística, producida entre 1930 y 2020, de la cual las materias de nuestra facultad hacen uso (destacamos [29], [45], [47], [38], [38], [58] y [60]). Asimismo hemos realizado una exhaustiva revision de los apuntes que los docentes de las materias del área comparten con sus alumnos. En particular, nos hemos nutrido de las notas del Prof Pablo Ferrari y de las elaboradas por la Prof Maria Eugenia Szretter y las del Prof Victor Yohai. Si bien este material no es de publico acceso, consideramos pertinente citarlo cuando expongamos ideas inspiradas o relacionadas con el mencionado material. A partir de esta incursión nos interesa proponer, por un lado, algunas discusiones en torno a las diferentes alternativas encontradas para pensar posibles organizaciones matemáticas de determinados asuntos hacia adentro de un aula. Consideraremos posibles potencialidades didácticas de estas maneras de organizar el discurso matemático, algunas de las cuales propondremos en el capítulo siguiente en diálogo con los problemas propuestos. Por otro lado, propondremos un recorrido por algunas demostraciones de los grandes teoremas que se tratan en esta tesis.

3.1. Algunas discusiones asuntos estocásticos que circulan en el aula

Un hecho en particular que atraviesa buena parte de esta sección está en relación con el rol que pueden cumplir los pasajes al límite para la definición, motivación o construcción de determinados objetos o propiedades que son relevantes para la probabilidad y/o la estadística.
Esta presumible conveniencia de la herramienta de pasaje al límite tiene su correlato en diversos contextos propios de la modelización de problemas mediante herramientas estocásticas: aproximación de fenómenos continuos mediante enfoques discretos, repetición “infinita"de experimentos, conceptualización de fenómenos continuos como límite (en algún sentido) de fenómenos discretos, se encuentran por doquier en los problemas de modelización, ya sean estos intra-matemáticos o extra-matemáticos. Consideramos que dar un lugar a ellos dentro de la construcción de teoría y de la construcción de conceptos puede ser una vía interesante, que exploraremos con diferentes enfoques en las siguientes líneas.

3.1.1. La probabilidad, la probabilidad condicional y la independencia

Esta tesis, como se ha referido, parte de una posición (que esperamos sustentar más profundamente en el siguiente capítulo) según la cual la construcción de conexiones entre las nociones probabilísticas y la experimentación, a partir de problemas que impliquen experimentar y/o simular situaciones aleatorias, es una vía interesante para dar los primeros pasos en la formación estocástica de nuestres estudiantes. Consideramos interesante que el estudio de una Teoría de Probabilidades esté en relación con los conocimientos y problemas en torno a la probabilidad que les estudiantes tienen, que una gran cantidad de veces están asociados al juego y a la experiencia. Según el momento de la formación en que nos encontremos y de la profundidad matemática que querramos imprimirle, será necesario construir precisa y formalmente una Teoría de Probabilidades que permita, en su lenguaje, modelar estos y otros problemas.

Esto nos expone a la tensión, no específica de ningún nivel educativo, entre la construcción de teoría matemática en el aula, y la inclusión en ella de los conocimientos que les estudiantes tienen y que producen al interactuar con problemas matemáticos. Observemos que, si nuestro esquema se limita a la espacios de probabilidad equiprobables, como lo son los asociados a juegos de azar (los cuales, como hemos visto en el capítulo anterior, están en el ínicio del desarrollo de la teoría de probabilidades como hoy la conocemos), una definición de probabilidad que puede ser razonable es $\frac{\text{casos favorables}}{\text{casos posibles}}$. Una gran cantidad de problemas y asuntos matemáticos de naturaleza estocástica pueden ser abordados con esta idea, posibilitando el desarrollo de intuiciones, heurísticas y razonamientos de mucha valía para la formación estocástica de les estudiantes. También permite un primer acercamiento a ciertas propiedades que cumple cualquier medida de probabilidad.

Ahora bien, fuera de los juegos de azar empieza a crujir la validez de esta definición, y surge una pregunta que nos interesa hacernos: ¿De qué manera se puede articular la noción frecuentista de la probabilidad con la teoría axiomática[1]? Es claro el correlato experimental de la primera, y su limitación formal (¿cómo abordar, desde este punto de vista, la sigma-aditividad?). La segunda, por su parte, puede -y, en cierta medida, debe- parecer artificial, e incluso al principio, disjunta con las cuestiones de ella que competen al abordaje de problemas concretos (¿qué sentido tiene, inicialmente, la formalidad de la noción de sigma-álgebra, considerando cómo su uso intensivo cobra particular relevancia recién en el tratamiento de los llamados *Tail events, o eventos de cola*?). De hecho, Kolmogorov introduce esta noción recién en el Capítulo 2 de su libro fundacional [43], mencionando que es "practicamente imposible dotar de sentido empírico a este axioma."(Kolmogorov, 1956, p. 14)

Desde ya, no hay respuesta única a esta tensión, que tiende a resolverse soportando una relativa coexistencia de ambos asuntos. Suele asumir una forma en la que el enfoque frecuentista funciona como una motivación y, a la vez, como un resultado provisorio e inspirador que podrá ser demostrado una vez presentada la teoría axiomatica. Pero a estos fines aparece uno de los resultados que mostramos en el capítulo anterior: que la existencia del límite de las frecuencias relativas puede sostenerse, apoyándose en lo demostrado por Jakob Bernoulli ([7]), siempre que:

- Se esté trabajando en un espacio equiprobable.
- No se esté trabajando en un espacio equiprobable, pero se tenga definida una noción de independencia que permita postular que la probabilidad de que un evento determinado ocurra k veces se corresponde con la función de probabilidad puntual de una variable aleatoria binomial.

También mencionaremos en este capítulo que, en términos modernos, podemos ver en [58] que Cantelli demuestra la Ley Fuerte de los Grandes Números para variables aleatorias Bernoulli, prosiguiendo y mejorando al demostración de Jakob. A partir de lo que acabamos de mencionar, nos interesa indagar en dos conceptos que también guardan una profunda relación con la modelización de problemas: la independencia y la probabilidad condicional.
Ambos conceptos tienen diferentes versiones cuando son mirados en espacios de probabilidad equiprobables, y cuando no. Por ejemplo, en un espacio de probabilidad equiprobable, se puede dar una definición de probabilidad condicional cercana al contexto de modelización de problemas:

$$P(A|B) = \frac{\text{Cantidad de eventos unitarios compatibles con } A \text{ y } B}{\text{Cantidad de eventos unitarios compatibles con } B}$$

que puede ser vista como una fórmula que restringe el espacio muestral al conjunto B Frecuentísticamente hablando, por su parte, una manera de motivar la definición de probabilidad condicional es la que presenta el Prof. Pablo Ferrari en sus notas, al considerar

$$\lim_{n\to\infty} \frac{\text{Cantidad de veces que ocurrió } A \text{ y } B}{\text{Cantidad de veces que ocurrió } B}$$

que resulta compatible con motivación de la probabilidad, pensada heurísticamente como límite de frecuencias relativas. Observemos que, frecuentísticamente, también se recupera la idea experimental

[1]Omitimos la discusión sobre el enfoque bayesiano

de restringir, del universo de observables, solo a aquellos que involucran la información de la ocurrencia del evento B. Tanto desde la óptica equiprobable como desde la frecuentista, resulta intuitivo proponer

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

siendo está la igualdad que define a la probabilidad condicional en los libros de texto, y que les autores suelen apoyar en algunas de las consideraciones que hemos mencionado líneas más arriba.
En todos los libros de probabilidad, por su parte, la definición de independencia es la misma y si bien puede motivarse desde la definición de probabilidad condicional, sortea la dificultad generada por eventos de probabilidad cero a la hora de condicionar, definiendo independencia en termino de la factorización de la probabilidad de la intersección de dos eventos. Específicamente, dos eventos A y B son independientes si

$$P(A \cap B) = P(A)P(B)$$

Ahora bien: los eventos no se llaman "interseccionalmente multiplicativos". Se llaman "independientes". Este nombre **convencional** tiene una raigambre en la modelización de fenómenos -tanto histórica como contemporáneamente-, y por ende cabe preguntarse por qué relaciones se pueden establecer en esta dirección.
Hay ciertos eventos que son claramente independientes (donde "claramente" quiere decir: seguramente muy pocas personas objeten, para estos eventos, su independencia). Por ejemplo: el resultado de tirar un dado, con respecto al resultado de tirar ese dado nuevamente. La definción matemática procura dar cuenta de la noción coloquial de este término. Como tantas veces, la matemática se apropia de palabras que forman parte del lenguaje cotidiano para darles una definición específica procurando reflejar su significado en la conversación corriente.

La idea de independencia puede estar apoyada, ciertamente, en que el espacio muestral asociado a la realización de dos experimentos puede expresarse como el producto de los espacio muestrales de cada uno de ellos. Es decir, si Ω_i, para $i = 1, 2$, denota el espacio muestral asociado al i-ésimo experimento, tenemos que $\overline{\Omega} = \Omega_1 \times \Omega_2$ permite registrar el resultado de cada par observado. Es decir, $\overline{\Omega}$ es el espacio muestral del experimento compuesto. Ahora bien, si cada espacio era equiprobable el experimento compuesta resulta nuevamente equiprobable en $\Omega_1 \times \Omega_2$. Por lo tanto, si queremos ahora calcular la probabilidad de que el resultado del experimento i pertenezca a A_i, podemos expresar esto en terminos del experimento compuesto. Luego, la probabilidad solicitada está dada por

$$\frac{\#(A_1 \times A_2)}{\#(\Omega_1 \times \Omega_2)} = \frac{\#A_1}{\#\Omega_1}\frac{\#A_2}{\#\Omega_2}.$$

De esta manera concluimos que el significado coloquial de independencia en escenarios equiprobables se condice con la definición de independencia dada: la probabilidad de que dos eventos independientes (en sentido coloquial) ocurran de manera simultanea está dada por el producto de las probabilidades de cada uno de ellos.

Esta idea, en el marco equiprobable, es coherente con la regla de multiplicatividad que se aplica para calcular el cardinal de ciertos conjuntos finitos, y que a nivel formal tiene un correlato concreto en cómo es la construcción de la medida de probabilidad en el espacio producto.
Otra idea de independencia suele estar sostenida la falta de influencia causal: la información de la ocurrencia de uno de los eventos no altera, de ningún modo, la probabilidad de ocurrencia del otro. Esta idea de independencia es un poco más robusta desde el punto de vista experimental, y tiene que ver con la reducción proporcional de las probabilidades. Los eventos "sacar un as de un mazo de cartas" y "sacar picas de un mazo de cartas" deben ser independientes, aunque causalmente uno implique una reducción del espacio muestral (si primero extraigo picas, estoy obligando a extraer ahora el as de picas). Esta idea, por su parte, es consistente con la caracterización

$$P(A|B) = P(A)$$

De hecho, a un nivel más formal, esta definición de independencia implica (si B no tiene probabilidad 0) la multiplicatividad de manera bastante directa, por lo que es un posible punto de partida. El asunto que aparece aquí es que, por un lado, no vale para eventos de probabilidad nula (consistentemente con que la probabilidad condicional no está definida en este caso), y por otro, que la idea de "falta de influencia causal" puede llevar a conclusiones erróneas, según cómo sea la medida de probabilidad y el espacio muestral en cuestión. No obstante, estas limitaciones formales tienen su contrapartida en las limitaciones intuitivas de la definición más general: si nos quedamos con ella, dejamos de poder describir y pensar numerosas situaciones donde las ideas de independencia, ya sean las asociadas a falta de influencia causal como a las de multiplicatividad del espacio muestral. Para variar, la tensión entre el alcance formal y el alcance de la modelización requiere una solución que no implica el predominio de uno sobre el otro, sino la articulación de ambos.

3.1.2. Las variables aleatorias

Entre los objetos estocásticos ineludibles se encuentran, sin duda, las variables aleatorias, y con especial mención aquellas llamadas "famosas". Una organización usual que reciben en los libros es, una vez definida la axiomática de Kolmogorov y realizado un trabajo en torno a ella y a problemas de probabilidad, probabilidad condicional e independencia, se definen las variables aleatorias. Una posible manera de introducir este tema es presentando a la variable aleatoria como una caracteristica asociada a cada posible resultado del experimento, expresada como como funciones del espacio muestral en $\mathbb{R}$. Es una *observable*, entendida como un valor que observamos o calculamos una vez obtenido el resultado del experimento.

En muchas situaciones la observable puede coincidir con el propio resultado del experimento. Por ejemplo: para el experimento "tirar un dado", es natural considerar la variable $X =$ "*número que sale en el dado*" y el cálculo de su función de probabilidad puntual no representa mayor dificultad sin siquiera pasar por el espacio muestral. Bishop [11], entre otros autores, evita hablar de espacios muestrales y se embarca de lleno en las variables aleatorias (y vectores) y el cálculo de probabilidad asociadas eventos de interés que son expresados en términos de las mismas: probabilidades puntuales, acumuladas, conjuntas, marginales.

El cálculo de probabilidades asociadas a variables aleatorias definidas en un espacio de probabilidad (Ω, F, P) requiere un manejo algebraico de nociones como imagen inversa y sus propiedades siendo que la función de probabilidad P está definida en el espacio (Ω, F) y es *transportada* para poder asignarle probabilidad a conjuntos *medibles* en $\mathbb{R}$ mediante la fórmula $P_X(A) = P(X \in A)$. Típicamente el rigor que se le de a esta problematica está asociado con la madurez matemática de los estudiantes. Estos son temas ampliamente estudiados en diferentes formaciones y adaptables en su formalismo, procurando no perder de vista la caracteristica primordial del concepto: estamos ante una observable asociada a un experimento incierto y necesitamos poder asginar probabilidades a los valores que esta observable pueda tomar. Con o sin espacio muestral, esta es una idea central que necesitamos poder transmitir.

Con las variables aleatorias famosas puede darse un fenómeno comparable: mientras que algunas pueden ser construidas con referencia a asuntos estocásticos concretos vinculados a los juegos de azar (la distribución binomial, bernoulli y geométrica son buenos ejemplos de esto), otras pueden ser motivadas o puestas en juego en otro tipo de situaciones. La incursión en la bibliografía permitió encontrar algunas referencias concretas que pueden permitir enfocar a algunas variables famosas en este sentido, ya sea como:

1. Variables cuya fórmula surge del planteo de problemas de modelización específicos, como es el caso de la variable aleatoria de Cauchy (intersección de un haz de luz proveniente de un punto dado, con una recta que no pasa por ese punto).

2. Variables que se deducen de experimentos con probabilidad condicional (deducción de la Beta como a posteriori de la binomial)

3. Variables que pueden ser vistas como procesos estocásticos unidimensionales a tiempo discreto, ya sea por refinamiento de experimentos discretos (la uniforme continua como límite de uniformes discretas, la binomial y geométrica como límite de extracciones de urna) o por discretización de un determinado experimento (la exponencial como límite de geométricas, la Poisson como límite de Bernoulli's)

Consideramos que estas posibilidades, en un aula, pueden ser fuente de discusiones y de interés para les estudiantes. Por ejemplo, que el límite de geométricas sea la exponencial, es una manera posible de abordar por qué se utiliza un modelo exponencial para la vida útil de lámparas de luz. Otro ejemplo es la sucesión de v.a. "cantidad de bolillas negras extraídas con reposición de una urna con N negras y B blancas", que es interesante para discutir con la asunción, promovida por Quetelet a mediados del siglo XIX, en torno a que todo fenómeno podía describirse, bien aproximado, por un modelo de urnas. No es soslayable, por su parte, que en la organización usual de los contenidos la idea de convergencia en distribución usualmente es posterior, incluso, a la idea de vector aleatorio. Conviene preguntarse, por otro lado, si estas consideraciones no permiten pensar en condiciones de posibilidad para la construcción de ese asunto matemático complejo justamente a partir del estudio de variables aleatorias famosas.

3.1.3. La esperanza, la varianza y la desigualdad de Bienaymé-Tchebychev

Antes de abocarnos a los grandes teoremas, se encuentra la noción de esperanza -y su noción solidaria, la varianza- que son objeto de abundante trabajo en el aula, y cuya multiplicidad de sentidos puede ser fuente de diversas discusiones.

Omitiremos una discusión pormenorizada sobre los diversos sentidos que puede tener la esperanza. Partiendo de la definición de esperanza para variables discretas, nos interesa problematizar una posible aproximación la definición de esperanza para variables continuas. De ahora en más, dada una variable aleatoria X, si esta es una variable discreta definiremos su función de probabilidad puntual como $p_X(k) = \mathbb{P}(X = k)$, para todo k tal que $P(X = k) \neq 0$. Si la función $f_X(x) = F'_X(x)$ -donde $F_X(x) = \mathbb{P}(X \leq x)$)- está bien definida, la llamaremos *densidad* de la variable X.
Usualmente, una manera de definir la esperanza de variables continuas es por extensión de la fórmula para variables discretas, reemplazando, en la definición de esperanza para discretas

$$\mathbb{E}(X) = \sum_{k \in R_X} p_X(k)k$$

el símbolo de sumatoria por el de integral, y la probabilidad puntual por la densidad (si existe) multiplicada por el diferencial dx, quedando como sigue:

$$\mathbb{E}(X) = \int_{\mathbb{R}} f_X(x)x\,dx$$

Este enfoque tiene, como todos, sus potencialidades y sus limitaciones: ¿en qué sentido este reemplazo simbólico tiene un correlato en la modelización de fenómenos? Es interesante el uso de la discretización que se realiza en el libro **A natural introduction into probability** (2003, [45]), que describimos a continuación, y que ofrece otra manera de aproximarnos a la definición.
Dada X variable aleatoria, si consideramos la discretización

$$X_n = \sum_{k \in \mathbb{Z}} \frac{k}{n} \mathbb{I}_{\{\frac{k}{n} < X \leq \frac{k+1}{n}\}}$$

se tiene que

$$\begin{aligned}
\mathbb{E}(X_n) &= \sum_{k \in \mathbb{Z}} \frac{k}{n} P\left(X_n = \frac{k}{n}\right) \\
&= \sum_{k = \in \mathbb{Z}} \frac{k}{n} P\left(\frac{k}{n} < X < \frac{k+1}{n}\right) \\
&= \sum_{k \in \mathbb{Z}} \frac{k}{n} \int_{k/n}^{(k+1)/n} f_X(x)dx \\
&= \int_{-\infty}^{+\infty} s_n(x)dx
\end{aligned}$$

donde

$$s_n(x) = \frac{k}{n} f_X(x) \quad \text{si} \quad \frac{k}{n} \leq x < \frac{k+1}{n}.$$

Observemos que, de la definición de $s_n(x)$, resulta que

$$|s_n(x) - xf(x)| \leq \frac{1}{n} f_X(x)$$

Por lo que:

$$\begin{aligned}
\left| E(X_n) - \int_{-\infty}^{\infty} xf(x)dx \right| &= \left| \int_{-\infty}^{\infty} (s_n(x) - xf(x))dx \right| \\
&\leq \int_{-\infty}^{\infty} |s_n(x) - xf(x)|\, dx \\
&\leq \frac{1}{n} \int_{-\infty}^{\infty} f_X(x)dx = \frac{1}{n},
\end{aligned}$$

de modo que $E(X_n) \to \int_{-\infty}^{\infty} f_X(x)x dx$. Consideramos que el resultado demostrado, que parte de discretizar un problema continuo, puede ser tomado como un buen disparador para la definición de la esperanza en el caso continuo[2].

Habiendo motivado la definición de esperanza en el caso continuo, nos interesa estudiar las relaciones entre la esperanza, la varianza y el desvío estándar. Más específicamente, queremos abordar maneras en que la matemática expresa la relación entre varianza/desvío y la dispersión de la variable aleatoria en torno a su esperanza, y qué sentidos posibles podemos recuperar con su trabajo en el aula.

En los cursos de Probabilidad y Estadística, uno de los resultados que cuantifican esta relación es la desigualdad de Markov, que suele ser necesaria para demostrar la desigualdad de Bienaymé-Thebychev, que se utiliza más extensivamente. Es interesante observar que la primer formulación en la historia de esta desigualdad es, en lenguaje moderno, la siguiente:

Proposición 3.1.1 *Bienaymé 1852, [18] Tchebychev 1867, [56]*
Sean $X_1, \ldots, X_n$ variables aleatorias independientes de rango finito con segundo momento finito, y sea $\alpha > 0$. Entonces,

$$P\left(\left| \sum_{k=1}^{n} X_i - \sum_{k=1}^{n} E(X_i) \right| > \alpha \sqrt{\sum_{k=1}^{n} Var(X_i)} \right) \leq \frac{1}{\alpha^2}$$

Observemos que, para una sola variable aleatoria, la formulación queda como sigue:

Proposición 3.1.2 *Bienaymé 1852, Tchebychev 1867*
Sean X variable aleatoria de rango finito con segundo momento finito, y sea $\alpha > 0$. Si σ_X es el desvío de X, entonces,

$$P(|X - E(X)| > \alpha\sigma_X) \leq \frac{1}{\alpha^2}.$$

[2]Observar que un razonamiento análogo permite extender la definición de esperanza para variables aleatorias con cualquier tipo de distribución, utilizando la integral de Riemman-Stieltjes. Una escritura posible de este argumento puede encontrarse en el apunte de Victor Yohai.

Esta desigualdad nos informa que el desvío estándar es una *unidad universal* para cuantificar la dispersión de una variable aleatoria respecto de su media. En este sentido, la desigualdad puede tener un sentido operatorio concreto: cuantificar cuánta masa (respect a P_X) concentran los valores que distan de la media (μ) más que una determinada cantidad (α) de desvíos estándar σ_X.
Por su parte, la formulación que aparece mayoritariamente en los libros de texto es la siguiente:

$$P\left(|X-E(X)|>\alpha\right) \leq \frac{Var(X)}{\alpha^2},$$

que por su parte permite cuantificar el hecho de que, a menor varianza, mayor cantidad de masa se encuentra en torno al valor esperado. Nuevamente, formulaciones matemáticas equivalentes comportan sentidos potencialmente diferentes para les estudiantes.

3.2. Los grandes teoremas

En esta sección, como adelantamos, enunciaremos diferentes versiones de los teoremas que esta tesis aborda. Tanto su formulación, como sus demostraciones, pueden ser elementos interesantes para considerar en las discusiones de aula.
Ahora bien: antes de encomendarnos a esta tarea, es interesante detenernos parcialmente a preguntarnos por qué sentidos pueden construirse para estos teoremas con relación a una mirada sobre problemas donde estos teoremas puedan jugar un rol para resolverlos y/o motivar su conjetura. Especialmente esto cobra relevancia cuando comparamos la ley débil y fuerte de los grandes números. ¿Cómo se interpreta la afirmación $P\Big(\{\omega\in\Omega/\overline{X_n(\omega)}\to E(X)\}\Big)=1$? Dejamos abierta esta pregunta, que nos interpela y cuya problematización puede ser interesante para futuras investigaciones.

De ahora en más, diremos que un proceso estocástico $\{X_n\}_{n\in N}$ verifica la ley de los grandes números si

$$\overline{X_n}-\mathbb{E}\left(\overline{X_n}\right)\to 0\,,\quad \text{siendo}\quad \overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i. \tag{3.1}$$

Diremos que la ley es débil o fuerte según si la convergencia sea en probabilidad o casi todo punto.

3.2.1. Ley débil de los Grandes Números

Como mencionamos en el capítulo anterior, el primer registro que consta de este resultado es el que sigue:

Proposición 3.2.1 (Jakob Bernoulli, 1730, [7], [58]) *Sean $\{X_n\}_{n\in N}$ variables aleatorias Bernoulli i.i.d. Entonces vale la ley débil de los grandes números.*

Omitiremos la demostración de este teorema por su extensión y complejidad técnica. De todas formas, puede leerse en detalle en el Capítulo 6 del libro de Uspensky, [58]. Recordemos que, si bien la demostración original asume que p es racional, la prueba puede extenderse al caso general. Nos parece pertinente mencionarla en primer lugar, no solo por ser la primera en la historia, sino porque apela solamente a herramientas de análisis matemático, junto a la asunción (sin demostración) de que, para un experimento con dos resultados posibles, cada uno con probabilidad p y $1-p$, la probabilidad de que haya k ocurrencias de uno y $n-k$ ocurrencias de otro esta dada por el coeficiente correspondiente del desarrollo del binomio $(p+(1-p))^n$. En las materias de nuestra facultad, usualmente se suele dar la siguiente versión de la ley débil:

Proposición 3.2.2 *Sea $\{X_n\}_{n\in N}$ variables aleatorias i.i.d con segundo momento finito. Entonces vale la ley débil de los grandes números.*

Asumiendo conocida la distribución de la suma de variables normales independientes puede ser interesante comenzar por la demostración en el caso normal puesto que, en este caso, hay formula cerrada para el cálculo de probabilidades, sin necesidad de invocar desigualdades. Más aún, este trabajo puede desarrollarse despues de una experimentación computacional, como se propone en el Problema 5 del Capítulo 4.

Si $X_i \sim N(\mu,\sigma^2)$, tenemos que, si $Z=\frac{\overline{X}_n-\mu}{\frac{\sigma}{\sqrt{n}}}$, entonces $Z\sim N(0,1)$ y luego

$$P\left(|\overline{X}_n-\mu|>\varepsilon\right)=P\left(\left|\frac{\overline{X}_n-\mu}{\frac{\sigma}{\sqrt{n}}}\right|>\frac{\varepsilon}{\frac{\sigma}{\sqrt{n}}}\right)=P\left(|Z|>\frac{\sqrt{n}\varepsilon}{\sigma}\right) \tag{3.2}$$

y esta última expresión converge a cero cuando $n\to\infty$. Queda así demostrado el caso normal, valiendonos del cálculo de probabilidades siendo que podemos caracterizar la distribución del promedio de variables normales iid.

Ahora bien, el enunciado nada dice sobre la distribución de las variables aleatorias. No tenemos información para calcular $P\left(|\overline{X}_n-\mu|>\varepsilon\right)$. Sin embargo, la desigualdad de Bienaymé-Tchebychev da una cota que permite concluir el resultado enunciado. Cabe mencionar que calcular probabilidades bajo normalidad y usar cotas universales permite abrir un debate interesante sobre las bondades y limitaciones de cada escenario. Conocer el modelo da lugar a resultados más precisos, mientras que las cotas permiten controlar casos más generales a expensas de perder eficiencia.

En este sentido, consideramos que es interesante haber realizado un trabajo en el aula con esta desigualdad, de modo que su uso pueda, en el contexto general de una demostración, apoyarse en una experiencia de trabajo. La desigualdad de Bienaymé-Tchebychev permite demostrar la ley débil en escenarios más generales, como muestra el siguiente resultado, presentado en el Teorema 5.6 de [38]

Proposición 3.2.3 *Sean $\{X_n\}_{n\in N}$ variables aleatorias no correlacionadas y supongamos que existe C tal que $Var(X_i)\leq C$, para todo i. Entonces vale la ley débil de los grandes números.*

En el libro Stochastics, de Georgii, se puede encontrar una demostración de la ley débil para variables iid (enunciado en el Teorema 5.7), asumiendo apenas primer momento finito.

Proposición 3.2.4 *Sea $\{X_n\}_{n\in N}$ una sucesión de variables aleatorias iid de primer momento finito. Entonces vale la ley débil de los grandes números.*

Antes de terminar esta sección, es interesante referirnos a una versión del teorema en otro escenario, cuyo detalle puede leerse en Uspensky:

Proposición 3.2.5 *Markov, en Uspensky (1937) [58]*
Sean $\{X_n\}_{n\in N}$ *variables aleatorias independientes. Si existen* $\delta > 0$ *y* M *de modo que, para todo* $n \in \mathbb{N}$ $E|X_n|^{1+\delta} \leq M$*, entonces vale la ley débil de los grandes números.*

Ley fuerte de los grandes números

Presentamos ahora una demostración de la ley fuerte para el caso de variables iid con segundo momento finito.

El siguiente resultado establece la validez del la ley fuerte para variables iid con primer momento finito.

Teorema 3.2.6 *Sea* $\{X_n\}_{n\in N}$ *una sucesión de variables aleatorias i i.i.d. con* $E(|X_i|) < \infty$*. Entonces vale la ley fuerte de los grandes números.*

Presentaremos la demostración para el caso en que las variables aleatorias tienen segundo momento finito. La demostración del caso general puede verse en el Teorema 2.4.2 de [29].

Para comenzar, utilizaremos la siguiente condición suficiente para la convergencia casi segura, que se deduce del lema de Borel Cantelli:

$$\text{Si} \quad \sum_{n=1}^{\infty} P\left(|Y_n - Y| > \varepsilon\right) < \infty \quad \forall \varepsilon > 0 \Rightarrow Y_n \to Y \quad \text{casi seguramente.} \tag{3.3}$$

Observemos que la formulación propuesta en (3.1) nos permite reducirnos al caso en que las variables X_i tienen esperanza 0.
Luego, de (3.3), alcanza ver que

$$\sum_{n=1}^{\infty} P\left(\left\{|\overline{X_n}| > \varepsilon\right\}\right) < \infty \quad \forall \varepsilon > 0$$

Observemos que, por Tchebychev, si $Var(x_1) = \sigma^2$, entonces

$$P\left(\left\{|\overline{X}_n| > \varepsilon\right\}\right) \leq \frac{1}{\varepsilon^2} V(\overline{X}_n) = \frac{\sigma^2}{\varepsilon^2 n}.$$

por lo que la subsucesión $\{\overline{X}_{n^2}\}_{n\in\mathbb{N}}$, verifica que

$$P\left(\left\{|\overline{X}_{n^2}| > \varepsilon\right\}\right) \leq \frac{\sigma^2}{\varepsilon^2 n^2},$$

y luego, $\{\overline{X_{n^2}}\}_{n\in\mathbb{N}}$ converge a 0 casi seguramente.

Usaremos los siguientes resultados:

- Si $0 \leq Z_n \leq Y_n$ y $Y_n \to 0$ casi seguramente, entonces $Z_n \to 0$ casi seguramente.

- $Z_n \to 0$ casi seguramente si y solo si $|Z_n| \to 0$ casi seguramente

- para cada $k \in \mathbb{N}$, existe un $n = n(k)$ que verifica $n^2 \leq k < (n+1)^2$

Entonces, podemos acotar:

$$0 \leq |\overline{X}_k| = \frac{1}{k}\left|\sum_{i=1}^{k}\overline{X}_i\right| = \frac{1}{k}\left|\sum_{i=1}^{n^2(k)}\overline{X}_i + \sum_{i=n^2(k)+1}^{k}\overline{X}_i\right| \tag{3.4}$$

$$\leq \frac{1}{n^2(k)}\left|\sum_{i=1}^{n^2(k)}\overline{X}_i\right| + \frac{1}{k}\left|\sum_{i=n^2(k)+1}^{k}\overline{X}_i\right| \tag{3.5}$$

y observando que

$$\frac{1}{n^2(k)}\left|\sum_{i=1}^{n^2(k)}X_i\right| = \left|\overline{X}_{n^2(k)}\right| \xrightarrow[k\to\infty]{c.s.} 0$$

Solo nos resta ver que

$$\frac{1}{k}\left|\sum_{i=n^2(k)+1}^{k}\overline{X}_i\right| \xrightarrow[k\to\infty]{c.s.} 0$$

Para ello, observando que

$$E\left(\sum_{i=n^2(k)+1}^{k}\overline{X}_i\right) = 0.$$

se tiene nuevamente con la desigualdad de Bienaymé-Tchebychev, que

$$P\left(\frac{1}{k}\left|\sum_{i=n^2(k)+1}^{k}\overline{X}_i\right| > \varepsilon\right) = P\left(\left|\sum_{i=n^2(k)+1}^{k}\overline{X}_i\right| > \varepsilon k\right) \leq \frac{1}{\varepsilon^2 k^2}V\left(\sum_{i=n^2(k)+1}^{k}\overline{X}_i\right) = \frac{1}{\varepsilon^2 k^2}\sigma^2(k-n^2(k)). \tag{3.6}$$

se tiene lo que afirmamos, puesto que

$$k - n^2(k) < (n(k)+1)^2 - n^2(k) = 2n(k)+1 \quad \Rightarrow \quad k - n^2(k) \leq 2n(k) \leq 2\sqrt{k}. \tag{3.7}$$

y finalmente

$$\sum_{k\geq 1} P\left(\frac{1}{k}\left|\sum_{i=n^2(k)+1}^{k}X_i\right| > \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2}\sum_{k\geq 1}\underbrace{\frac{2\sqrt{k}}{k^2}}_{\frac{2}{k^{3/2}}} < \infty. \tag{3.8}$$

La demostración que acabamos de hacer vale también para demostrar la ley fuerte bajo las hipotesis de la Proposición 3.2.3, como demuestra Georgii en el Teorema 5.16

Por último, nos interesa estudiar el comportamiento del promedio cuando las variables son iid pero sin primer momento finito, tanto débilmente como fuertemente. En esta dirección analizaremos el ejemplo que propone [29]:

Proposición 3.2.7 (Teorema 2.4.5) *Sean $\{X_n\}_{n\in\mathbb{N}}$ variables i.i.d con $EX_i^+ = \infty$ y $EX_i^- < \infty$. Si $S_n = X_1 + \cdots + X_n$ tentonces $S_n/n \to \infty$ casi seguramente.*

Demostración 3.2.8 *Sea un $M > 0$ y $X_i^M = \text{mín}\{X_i, M\}$. Las variables X_i^M son i.i.d. con $\mathbb{E}\left|X_i^M\right| < \infty$, luego consideremos $S_n^M = X_1^M + \cdots + X_n^M$. Por LGN, se tiene que $S_n^M/n \to EX_i^M$. Como $X_i \geq X_i^M$, se sigue que*

$$\liminf_{n\to\infty} S_n/n \geq \lim_{n\to\infty} S_n^M/n = EX_i^M$$

Por lo que el teorema de convergencia monótona afirma que $E\left(X_i^M\right)^+ \to EX_i^+ = \infty$ cuando $M \to \infty$. Luego, $EX_i^M = E\left(X_i^M\right)^+ - E\left(X_i^M\right)^- \to \infty$, y se tiene $\liminf_{n\to\infty} S_n/n \geq \infty$, lo que implica el resultado en cuestión.

3.2.2. Frecuencia relativa y probabilidad

Antes de adentrarnos en el teorema central del límite, nos interesa poner en diálogo las referencias históricas recuperadas en el capítulo anterior, y lo que mostramos en este capítulo asociado a la demostración de la ley de los grandes números.

Kolmogorov dota a la matemática de una teoria de la probabilidad, postulando una colección de axiomas, que llevan directamente a conclusiones en acuerdo con la experimentación práctica: la frecuencia relativa se estabiliza en un valor que, coloquialmente, denominamos probabilidad. En particular, si $(X_i)_{i\geq 1}$ son variables aleatorias iid, tenemos que $I_{X_i\in A}$ tambien son iid, con distribución Bernoulli de parámetro $p = P(X \in A)$, siendo $X \sim X_i$. Invocando la Ley de los Grander Números, concluimos que

$$\frac{1}{n}\sum_{i=1}^{n} I_{X_i\in A} \longrightarrow P(X \in A).$$

Es decir, la popular idea de que la *probabilidad* representa la frecuencia relativa con la que observamos cierto evento puede ser demostrada partiendo de los axiomas postulados. No obstante, la demostración realizada por Jakob Bernoulli alrededor del 1730, como mencionamos, sólo depende de la noción de independencia de dos eventos asumiendo que p es racional, siendo que trabajaba con *espacios equiprobables*, en términos modernos. Es decir que, en el marco axiomático de Kolmogorov, dicha demostración ya es matemáticamente válida sin necesidad de desarrollar un mayor despliegue teórico. Más aún, su demostración es valida incluso para todo p en el intervalo $(0, 1)$.

3.2.3. Teorema Central del Límite

Comenzamos esta sección recordando el enunciado del Teorema Central del Límite, recordando que si $\{X_n\}_{n\in N}$ es una sucesión de variables aleatorias, diremos que

$$X_n \xrightarrow{D} X$$

si las funciones de probabilidad acumulada $F_{X_n}(x)$ verifican

$$F_{X_n}(x) \underset{n\to\infty}{\longrightarrow} F_X$$

para todo x punto de continuidad de $F_X(x)$.

Teorema 3.2.9 *Sea $\{X_n\}_{n\in N}$ una sucesión de variables aleatorias i.i.d con segundo momento finito. Entonces, vale que*

$$\frac{\sqrt{n}\left(\overline{X_n}-\mu\right)}{\sigma} \xrightarrow{D} Z \sim N(0,1), \tag{3.9}$$

siendo $\mu = E(X_1)$, y $\sigma^2 = Var(X_1)$.

En esta sección, nos interesa hacer una revisión de las diferentes demostraciones de este teorema que se abordan en nuestra facultad. Nos interesa indagar en qué herramientas o técnicas se emplean para hacerlo, qué resultados teóricos pueden ser abordados y cuáles son invocados sin demostración. Si bien no seremos exhaustivos con todas las demostraciones, sí nos interesa dejar explícitos los asuntos antes mencionados, a fin de que a la hora de tomar decisiones en torno al abordaje de la demostración de este teorema, estas notas puedan ser de utilidad.

1. La primer demostración que nos interesa considerar puede encontrarse en Wasserman (2010, [60], Apéndice 5.7) es la que se da frecuentemente en la materia "Probabilidades y Estadística" para Computación, que utiliza (sin demostrar) el siguiente resultado, presentado en el Lema 5.20 de [60].

   Proposición 3.2.10 *Sea $\{Z_n\}_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias, Z es otra variable aleatoria. Sean ψ_n y ψ las funciones generadoras de momentos de Z_n y Z respectivamente. Si $\psi_n(t)$ converge a $\psi(t)$ para todo t en un entorno del origen, entonces Z_n converge a Z en distribución.*

   La demostración de este hecho requiere del uso de técnicas bastante poco elementales, que apelan a concebir a la FGM como un caso particular de la Transformada de Laplace, y estudiar sus propiedades.

2. Otra posible manera de demostrar el TCL es mediante la utilización de funciones características. Dada una variable aleatoria X, su función característica es la función $\phi_X(t) := \mathbb{E}(e^{itX})$
   En lo que sigue, enunciaremos los principales resultados con los que [23] demuestra este resultado. En este libro, se utiliza extensivamente la correspondencia entre una medida de probabilidad μ, y una función de distribución F. En lo que sigue, adaptaremos el abordaje realizado en dicho libro en términos de funciones de distribución.

   Proposición 3.2.11 (Teorema 6.1.4 del libro) *La función característica de la suma de finitas variables aleatorias independientes es el producto de sus respectivas funciones características.*

   Luego, utilizando fuertemente resultados de teoría de la medida, demuestra que

Proposición 3.2.12 (Teorema 6.2.1 del libro: la fórmula de inversión) *Sea X una variable aleatoria, F_X su función de probabilidad acumulada, ϕ_X su función característica, y $x_1 < x_2$ puntos de continuidad de F_X. Entonces*

$$F_X(x_2) - F_X(x_1) \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx_1} - e^{-itx_2}}{it} \phi_X(t) dt.$$

Esta proposición permite probar el siguiente hecho utilizando que toda función de distribución tiene a lo sumo numerables discontinuidades:

Proposición 3.2.13 (Teorema 6.2.2 del libro: la unicidad de la función característica) *Sean X e Y dos variables aleatorias tales que $\phi_X = \phi_Y$, entonces $F_X = F_Y$*

Con esto en mente, el autor requiere del siguiente resultado extra:

Teorema 3.2.14 (Teorema 6.3.2 del libro, conocido como el Teorema de Levy-Cramer.) *Sean $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias y sean $\{\phi_n\}_{n\in\mathbb{N}}$ sus respectivas funciones características. Supongamos que existe una ϕ_∞ continua en 0 tal que, para todo $x \in \mathbb{R}$,*

$$\phi_n(x) \to \phi_\infty(x)$$

Entonces:

- *$\{X_n\}_{n\in\mathbb{N}}$ converge en distribución a una X_∞*
- *la función característica de X_∞ es ϕ_∞*

Finalmente, el resultado que queda demostrar es el que permite dar una expansión como serie de potencias para las funciones características. Para esto, es preciso relacionar la derivabilidad de la función característica con los momentos de la variable aleatoria asociada. Concretamente:

Proposición 3.2.15 *Sea X una variable aleatoria, y $k \in \mathbb{N}$. Si $|X|$ tiene k-ésimo momento finito, entonces la k-ésima derivada $\phi_X^{(k)}$ resulta continua en un entorno de $t = 0$, acotada, y su fórmula es $\phi_X^{(k)}(x) = \int_{\mathbb{R}} (it)^k e^{itx} dF(t)$.*
Recíprocamente, si $\phi_X^{(k)}$ resulta continua en un entorno de $t = 0$, entonces $|X|$ tiene k-ésimo momento finito.

Corolario de este teorema, utilizando la existencia del desarrollo de Taylor en un entorno de $t = 0$ para funciones k veces derivables en $t = 0$, resulta que Si $|X|$ tiene k-ésimo momento finito, entonces

$$\phi_X(x) = \sum_{j=0}^{k} \frac{i^j}{j!} \mathbb{E}\left(X^j\right) x^j + o\left(|t|^k\right)$$

Con estos resultados probados, la LGN se demuestra (teorema 6.4.3) usando que, si $S_n = \sum_{i=1}^n X_i$, entonces, y gracias al teorema 6.1.2 del libro, para cada x:

$$E\left(e^{ix(S_n/n)}\right) = E\left(e^{i(x/n)S_n}\right) = \left[\phi\left(\frac{x}{n}\right)\right]^n = \left(1 + i\mu\frac{x}{n} + o\left(\frac{x}{n}\right)\right)^n \to e^{i\mu x}$$

Como la función característica de la variable aleatoria $X = \mu$ es exactamente $\phi_X(x) = e^{i\mu x}$, se tiene que $\overline{X_n}$ converge en distribución a dicha variable, lo cual implica a su vez que lo hace en probabilidad. Conviene mencionar que esta demostración no requiere segundo momento finito, por lo cual es interesante considerar el aporte que este constructo matemático aporta para lograr demostrar este resultado.
Análogamente el TCL, planteando (Teorema 6.4.4) que, si la esperanza de X_1 e s0, y tiene varianza finita σ^2, entonces:

$$E\left(e^{ix\left(\frac{S_n}{\sigma\sqrt{n}}\right)}\right) = \left[\phi\left(\frac{x}{\sigma\sqrt{n}}\right)\right]^n$$
$$= \left\{1 + \frac{i^2\sigma^2}{2}\left(\frac{x}{\sigma\sqrt{n}}\right)^2 + o\left(\frac{|x|}{\sigma\sqrt{n}}\right)^2\right\}^n$$
$$= \left\{1 - \frac{x^2}{2n} + o\left(\frac{x^2}{n}\right)\right\}^n \to e^{-x^2/2}.$$

lo cual demuestra el teorema, ya que la función característica de una variable $Z \sim N(0,1)$ es justamente $\phi_Z(x) = e^{-x^2/2}$

3. Otra manera de demostrar este teorema es relacionando la convergencia en distribución con la convergencia de las esperanzas de las variables $g(X_n)$, para toda g continua, y con todas sus derivadas acotadas y uniformemente continuas hasta su 2do orden, como realiza [38]. Con este resultado, la demostración puede realizarse para variables con segundo momento finito -con menos hipótesis el resultado no es cierto-.

Primero necesitamos los siguientes dos resultados (demostramos el primero, el segundo citamos dónde encontrar la demostración):

Teorema 3.2.16 (Teorema de Convergencia Acotada) *Si existe M tal que para todo n, $|X_n|$ es una variable con rango en $[-M, M]$ y X_n tiende a X en probabilidad, entonces $E(X_n)$ tiende a $E(X)$.*

Demostración 3.2.17 *Para realizar esta demostración, alcanza con observar que*

$$|X_n - X| \leqslant \varepsilon I_{\{|X_n - X| \leqslant \varepsilon\}} + 2MI_{\{|X_n - X| > \varepsilon\}}$$

Por lo que

$$E|X_n - X| \leqslant \varepsilon P(|X_n - X| \leqslant \varepsilon) + 2MP(|X_n - X| > \varepsilon) \leqslant \varepsilon + 2MP(|X_n - X| > \varepsilon) \longrightarrow \epsilon \text{ cuando } n \to \infty$$

Teorema 3.2.18 (Teorema de Skorohod) *Sean X y X_n, $n \geq 1$, variables aleatorias con función de distribución acumulada F y F_n, respectivamente. Si $X_n \xrightarrow{D} X$, entonces existe un espacio de probabilidad $(\tilde{\Omega}, \tilde{F}, \tilde{P})$,* *una secuencia de variables aleatorias $(Y_n)_{n\geq 1}$ y otra variable Y, todas definidas en $(\tilde{\Omega}, \tilde{F}, \tilde{P})$, tales que (i) $Y_n \sim F_n$ y $Y \sim F$ y (ii) $Y_n \to Y$ casi seguramente.*

Observación 3.2.19 *De hecho, se puede tomar $Y_n = F_n^{-1}(U)$, $Y = F^{-1}(U)$, con $U \sim U(0,1)$ como se demuestra en el Teorema 25.6 de [10].*

Con dichos resultados a la mano, examinaremos la demostración del siguiente teorema.

Teorema 3.2.20 *Son equivalentes:*
a) $X_n \xrightarrow{D} X$
b) $E[g(X_n)] \to E[g(X)]$ para toda g función continua y acotada.

a) $\Rightarrow$ b) Sabemos $X_n \overset{D}{\to} X$. Sea g continua y acotada. Por el Teorema de Skorokhod, $\exists(\Omega, F, P)$ espacio de probabidad, y variables Y_n, Y definidas en allí tales que $Y_n \sim X_n, Y \sim X$, y Y_n converge casi seguramente a Y, lo cual en particular implica que Y_n converge en probabilidad a Y. Como g es continua, resulta que $g(Y_n)$ converge en probabilidad a $g(Y)$. Y como g es acotada, del Teorema de Convergencia Acotada resulta que $E(g(Y_n)) \to E(g(Y))$, y luego, como

$$E(g(Y_n)) = E(g(X_n)) \text{ y } E(g(Y)) = E(g(X))$$

queda probada la ida.

b) $\Rightarrow$ a) Sabemos que $\forall g : \mathbb{R} \to \mathbb{R}$ continua y acotada vale que $E(g(X_n)) \to E(g(X))$. Queremos ver que $X_n \overset{D}{\to} X$, i.e. $F_{X_n}(t) \to F_X(t) \quad \forall t$ punto de continuidad DE F_X. Escribiremos $F_X(t) = E\left(I_{(-\infty,t]}(X)\right)$. Dado $\delta > 0$, sea $g : \mathbb{R} \to \mathbb{R}$ continua y acotada tal que:

$$I_{(-\infty,t]}(x) \leqslant g(x) \leqslant I_{(-\infty,t+\delta]}(x). \tag{3.10}$$

Entonces, $\quad I_{(-\infty,t]}(X_n) \leqslant g(X_n)$ y tambien $g(X) \leqslant I_{(-\infty,t+\delta]}(X)$

Tomando esperanza, tenemos

$$F_{X_n}(t) = P(X_n \leqslant t) \leqslant E(g(X_n)) \quad y \quad E(g(X)) \leqslant F_X(t+\delta)$$

Tomamos límite superior obtenemos, usando la hipotesis:

$$\limsup_{n\to\infty} F_{X_n}(t) \leq \limsup_{n\to\infty} E(g(X_n)) = E(g(X)) \leqslant F_X(t+\delta)$$

Y luego,

$$\limsup_{n\to\infty} F_{X_n}(t) \leq F_x(t) \tag{3.11}$$

$\limsup_{n\to\infty} F_{X_n}(t) \leq F_x(t) Seaahorahcontinuayacotada, talque$

$$I_{(-\infty,t-\delta]}(x) \leqslant h(x) \leqslant I(-\infty,t](x) \Rightarrow I_{(-\infty,t-\delta)}(x) \leqslant h(x)$$

y

$$h(x_n) \leqslant I_{(-\infty,t]}(x_n)$$

Tomando esperanza, resulta

$$\underbrace{P(X \leq t-\delta)}_{F_X(t-\delta)} \leqslant E(h(X)) \text{ y } E\left(h\left(X_n\right)\right) \leqslant \underbrace{P\left(X_n \leqslant t\right)}_{F_{X_n}(t)}$$

Entonces

$$F_X(t-\delta) \leqslant E(h(X)) = \liminf_{n\to\infty} E\left(h\left(X_n\right)\right) \leq \liminf_{n\to\infty} F_{X_n}(t)$$

y luego

$$F_X(t-\delta) \leqslant E(h(X)) \leqslant \liminf_{n\to\infty} F_{X_n}(t)$$

Ahora, tomando $\delta_{\to 0}$ y usando que F_X es continua en t, tenemos:

$$F_X(t) \leq \liminf_{n\to\infty} F_{X_n}(t) \tag{3.12}$$

Combinando las desigualdades presentadas en (3.11) y (3.12), tenemos:

$$F_x(t) \leq \liminf F_{X_n}(t) \leqslant \limsup F_{X_n}(t) \leqslant F_X(t) \quad \square$$

Observación 3.2.21 *Observemos que la función g encontrada en* (3.10) *puede tomarse con k derivadas acotadas y continuas (y luego uniformemente continuas). Esto nos permite, en el teorema anterior, sustituir el ítem b) por el siguiente ítem b'):*

$$b')E\left[g\left(X_n\right)\right] \to E[g(X)] \tag{3.13}$$

para toda g función continua, acotada, k veces derivable, y con todas sus derivadas uniformemente continuas y acotadas.

Con este resultado, [38] deduce el TCL para variables con segundo momento finito, a partir de considerar que si $\{X_n\}_{n\in\mathbb{N}}$ verifican $E(X_1)=\mu$, $V(X_1)=\sigma^2$, las variables auxiliares $\tilde{X}_k=(X_k-\mu)/\sigma$ permiten reducir el problema al caso en el que $\mu=0$, $\sigma^2=1$.
En este nuevo escenario, consideramos $\{X_n\}_{n\in\mathbb{N}}$ variables iid con $E(X_1)=0$, $V(X_1)=1$, y considerando las variables $S_n=\sum_{i=1}^n X_i$, y la colección de variables $\{Y_n\}_{n\in\mathbb{N}}$ tales que $Y_i \sim N(0,1)$ para todo $i\in\mathbb{N}$ y son independientes de las X_i, el autor demuestra que, para toda f satisfaciendo la condición estipulada en (3.2.21) *para $k=2$*

$$E\left(f\left(\frac{S_n}{\sqrt{n}}\right) \to E\left(f\left(\frac{T_n}{\sqrt{n}}\right)\right)\right. \quad \text{, donde} \quad T_n=\sum_{i=1}^n Y_i \sim N(0,1)$$

Para hacerlo, el truco clave es considerar a

$$E\left(f\left(\frac{S_n}{\sqrt{n}}\right)-f\left(\frac{T_n}{\sqrt{n}}\right)\right)$$

como una suma telescópica, y ver que tiende a 0. *Para hacerlo, considera* $X_{j,n} = \frac{X_j}{\sqrt{n}}$, $Y_{j,n} = \frac{Y_j}{\sqrt{n}}$, $W_{i,n} = \sum_{j=1}^{i-1} Y_{j,n} + \sum_{j=i+1}^{n} X_{j,n}$, *y luego, denominando*

$$f\left(\frac{S_n}{\sqrt{n}}\right) - f\left(\frac{T_n}{\sqrt{n}}\right) = \sum_{i=1}^{n}\left[f\left(W_{i,n} + X_{i,n}\right) - f\left(W_{i,n} + Y_{i,n}\right)\right],$$

ya que $W_{i,n} + X_{i,n} = W_{i-1,n} + Y_{i-1,n}$ *para todo* $1 < i \leq n$. *Luego, la demostración se reduce a hacer el desarrollo de Taylor correspondiente a* $f\left(W_{i,n} + X_{i,n}\right)$ *y* $f\left(W_{i,n} + Y_{i,n}\right)$ *usando 3.2.21.*

Observación 3.2.22 *Conviene, a los fines de terminar esta sección, mencionar que la hipótesis de segundo momento finito no puede ser relajada en el TCL.*

Antes de terminar este capitulo, queremos presentar la siguiente generalización del TCL debida a Lindeberg, en [44]

Proposición 3.2.23 *Lindeberg, 1922, teorema 7.2.1 de [23]*
Supongamos que tenemos una sucesión $X_{n,k}$ *de variables aleatorias de esperanza nula y varianza* $\sigma^2_{n,k}$, *de modo que* $\sum_{k=1}^{n} \sigma^2_{n,k} = \tau^2_n < \infty$. *Si* $S_n = \sum_{k=1}^{n} X_{n,k}$. *Supongamos que*

$$\text{(Condición de Lindeberg):} \quad L_n(\epsilon) := \frac{1}{\tau_n^2}\sum_{k=1}^{n} E\left(X^2_{n,k}\mathbb{I}_{\{|X_{n,k}|\geq \epsilon\tau_n\}}\right) \rightarrow_{n\rightarrow\infty} 0 \tag{3.14}$$

Entonces

$$P\left(\frac{S_n}{\tau_n} \leq a\right) \rightarrow \phi(a) \quad \forall\ a \in \mathbb{R}$$

donde ϕ *es la función de probabilidad acumulada de una variable con distribución* $N(0,1)$.

Observación 3.2.24 *El caso en el que, para todo* k, $X_{n,k} = X_n$ *y son i.i.d, se verifica la condición de Lindeberg, ya que*

$$L_n(\epsilon) := \frac{1}{n\sigma^2}\sum_{k=1}^{n} E\left(X^2_n\mathbb{I}_{\{|X_n|\geq n\epsilon\sigma\}}\right) = \frac{1}{\sigma^2}E\left(X^2_n\mathbb{I}_{\{|X_n|\geq n\epsilon\sigma\}}\right) \rightarrow_{n\rightarrow\infty} 0$$

por lo que es un corolario de esta demostración el Teorema Central del Límite en la versión que formulamos.

Con esto damos por terminado este capítulo, en el que hemos propuesto algunas de las posibles discusiones en torno a posibilidades para la organización de un discurso matemático, incluyendo definiciones, propiedades, proposiciones y teoremas con la intención de llevarlo a las aulas. Consideramos que estas discusiones, apoyadas en un relevamiento bibliográfico como el que describimos, pueden constituir un aporte y un punto de partida para la elaboración de otras maneras de enfocar la enseñanza y el aprendizaje de conceptos de probabilidad y estadística. En el siguiente capítulo -dedicado a la propuesta didáctica que elaboramos- dar algunas claves para problematizar y profundizar en torno a la circulación y construcción del discurso matemático en el aula.

Capítulo 4

La propuesta didáctica

En este capítulo, estudiaremos una propuesta didáctica diseñada con la intención de abordar, en aulas de nivel superior, conceptos de la Probabilidad y la Estadística, entre ellos y especialmente la Ley Débil de los Grandes Números y el Teorema Central del Límite. Para realizar este estudio, y para producir esta secuencia, se pusieron en juego posiciones y conocimientos tanto del ámbito matemático como didáctico. Como toda propuesta, esta se encuentra producida desde un posicionamiento que repercute en las decisiones y recortes que forman parte de ella.

A fin de estructurar el desarrollo de este capítulo, lo dividiremos en cuatro secciones.
En la primera, de carácter más bien general en torno a nuestro posicionamiento didáctico consideraremos aportes de grandes teorías didácticas -que a su vez funcionan como un marco general de esta tesis-. Nos concentraremos esencialmente en tres cuestiones:

- Nuestra concepción sobre la clase de matemática y la circulación de conocimientos en un aula.
- El rol de los problemas y las interacciones en el aula en la construcción de conocimientos por parte de les estudiantes.
- Una mirada sobre las características del trabajo matemático de les estudiantes.

En la siguiente sección, nos interesa profundizar en cómo estos elementos, junto con aportes específicos en torno a la enseñanza-aprendizaje de la probabilidad y la estadística tanto a nivel nacional como internacional, se ponen en juego a propósito del diseño de la propuesta. En particular, desarrollaremos:

- La noción de modelización matemática (MM) enmarcada en trayectorias que aborden cuestiones de naturaleza estocástica.
- La construcción de teoría probabilística y el rol de la experimentación, ya sea por medios analógicos o vía simulación.
- El posible rol de la inferencia estadística en la construcción de conocimientos de naturaleza estocástica.

En consonancia con una amplia literatura en torno a la didáctica de la matemática asociada a procesos de enseñanza-aprendizaje de conceptos tanto de la probabilidad y la estadística (Batanero (2005) [4], Batanero y Borocvnik (2016) [5], Batanero, Godino y Cañízares (1987) [6], Borovcnik (2014) [13], Coronado, García-García, Arredondo y Araya Naveas (2022) [24], Esteley y Magallanes (2022a y b) [31] y [32], Gal (2005) [35], Tauber (2001) [55], Pfannkuch y Wild (1999) [61]), utilizaremos el término *estocástico* para referirnos a conceptos, relaciones o asuntos matemáticos modelizables articuladamente desde el punto de vista probabilístico o estadístico, con un interés por aquellos que puedan ser vistos simultáneamente desde ambos marcos.
La tercer sección aborda nuestra propuesta didáctica, y con relación a ella explicitaremos:

- Decisiones en torno a la organización de la clase y del trabajo autónomo de les estudiantes.
- Aspectos en torno a los asuntos que la propuesta aborda y la organización de la misma.
- Análisis didáctico-matemático de la propuesta en su totalidad.

La última sección es una recapitulación que incluye algunas preguntas que quedan pendientes de ser estudiadas.

4.1. Elementos teóricos de la Didáctica de la Matemática

La clase de matemática, desde nuestra perspectiva, será concebida como una comunidad de producción, en el sentido de que estudiantes y docentes, a propósito de un proyecto educativo conjunto, producen conocimientos tanto individual como colectivamente. Esta producción tiene cabida en un tiempo y lugar específicos, siguiendo los modos de hacer (necesariamente dinámicos) de esa comunidad. Estos elementos conforman una marca que distingue el conocimiento producido por dicha comunidad de otros conocimientos producidos por otras comunidades en otros momentos.
El proyecto educativo que convoca y reúne a docentes y estudiantes tiene, como una de sus finalidades, el aprendizaje por parte de les estudiantes de una serie de asuntos matemáticos ya existentes en la comunidad científica. La relación entre estos asuntos matemáticos y los conocimientos que les estudiantes producen en el aula es compleja. Nos resulta fértil, para comenzar a conceptualizar esta relación, la distinción que Guy Brousseau (Brousseau y Centeno, 1991, [16]) realiza entre "conocimiento" y "saber":

> "los conocimientos son los medios transmisibles (por imitación, iniciación, comunicación, etc.) pero no necesariamente explicitables, de controlar una situación y de obtener de ella un cierto resultado conforme a una expectativa y a una exigencia social. El saber es el producto cultural de una institución que tiene por objetivo identificar, analizar y organizar los conocimientos a fin de facilitar su comunicación" (citado por Bloch, I; 1999, P.7)." [12]

Esta distinción entre saber y conocimiento ha tenido una gran pregnancia en la didáctica de la matemática y la concebimos como central, toda vez que permite complejizar la relación existente entre lo que se aprende y lo que se pretende enseñar, en la que lo primero ya no es una parte de lo segundo, sino que son asuntos de naturaleza necesariamente diferente. Desde nuestro punto de vista, nos enfocaremos principalmente en el conocimiento de cada sujeto en un momento dado: ¿Qué situaciones le permite resolver, y cómo? ¿Cuál es su relación con otros conocimientos?

Guy Brousseau, nació en 1933 en Francia, y desarrolló, hacia principios de los años 70, la Teoría de Situaciones Didácticas (TSD), una de las Teorías de la Didáctica de la Matemática más importantes a nivel mundial. Originariamente orientada al ámbito escolar de nivel primario e inicial, sus aportes y desarrollos alcanzan a todos los niveles educativos. Una gran cantidad de Teorías Didácticas se apoyan o referencian en la TSD, entre las cuales se encuentran la Teoría Antropológica de lo Didáctico (TAD) de Yves Chevallard y el Enfoque Ontosemiótico de Juan D. Godino. A la primera nos referiremos en esta sección, mientras que el segundo predomina en la producción nacional e internacional en torno a la enseñanza-aprendizaje de la Probabilidad y la Estadística. Mención especial también requiere la Ingeniería Didáctica (ID) (Artigue, 1988, [3]) a la que también nos referiremos y que, en calidad de Metodología de Investigación, ha sido crucial para el desarrollo y difusión de la TSD.

Sadovsky (2005, [48]) refiere que para la TSD el conocimiento de les estudiantes no es fruto de una reproducción, sino de una construcción que tiene lugar a propósito de la interacción de les estudiantes con un medio resistente que, para la TSD

> "incluye tanto una problemática matemática inicial que el sujeto enfrenta, como un conjunto de relaciones, esencialmente también matemáticas, que se van modificando a medida que el sujeto produce conocimientos en el transcurso de la situación [didáctica]"(P.3) [48]

Esta situación problemática es propuesta por le docente a les estudiantes, para que estes autónomamente se propongan resolverla. La TSD concibe que este proceso no está determinado por la entrega de un enunciado escrito a les estudiantes, sino que requiere una o más instancias en la que docentes y estudiantes consensúan cuál es la tarea a realizar. Mejor lo refiere Patricia Sadovsky (op cit):

> "No basta "comunicar" un problema a un alumno para que ese problema se convierta en su problema y se sienta el único responsable de resolverlo. Tampoco basta que el alumno acepte esa responsabilidad para que el problema que resuelva sea un problema "universal" libre de presupuestos didácticos. Denominamos "devolución" a la actividad mediante la cual el docente intenta alcanzar ambos resultados" (P.13) [48]

Una vez enfrentades a la situación problemática, les estudiantes deberán comprometer su sistema de conocimientos realizando acciones para resolverla. No obstante, esta situación no estará diseñada para que los conocimientos ya construidos de les estudiantes funcionen óptimamente en su resolución: el medio, compuesto entonces por la situación problemática y los conocimientos puestos en juego, generará unas resistencias que les estudiante deberán superar a partir de adaptar, modificar y objetar sus acciones. En este sentido, le docente desempeña un rol particular, tal como lo refiere Brousseau (1988, [17]):

> "el trabajo del docente consiste pues, en proponer al alumno una situación de aprendizaje para que produzca sus conocimientos como respuesta personal a una pregunta y los haga

> funcionar o los modifique como respuesta a las exigencias del medio y **no a un deseo del maestro**. Hay una gran diferencia entre adaptarse a un problema que plantea el medio, insoslayable, y adaptarse al deseo del docente. (...) Es el sujeto quien se ubica en posición de interpretar los resultados de sus acciones buscando analizar si las decisiones tomadas se encaminan a su finalidad (la resolución del problema)." (P.62, la negrita es nuestra) [17]

Esta etapa de trabajo en la que les estudiantes esencialmente interactúan con el medio, mientras que le docente se abstiene de proporcionar los conocimientos que la situación busca movilizar, se conoce en la TSD como fase adidáctica, y es uno de los aportes centrales de esta teoría.

Ahora bien, en relación al trabajo matemático que realizan les estudiantes, nos apoyaremos en la Teoría Antropológica de lo Didáctico. Esta teoría tiene como principal antecedente la publicación de "La Transposición Didáctica: del saber sabio al saber enseñado" de Yves Chevallard (1985, [21]), texto que marcó un antes y un después en la Didáctica de la Matemática (y en la didáctica en general).
Para la TAD, la realización de toda tarea (por ejemplo, resolver una ecuación de segundo grado) en una determinada institución requiere, necesariamente, de la disponibilidad de una técnica para realizarla, de un discurso cuyo fin es la validación de la técnica (en términos de la TAD, un discurso tecnológico), y de una teoría en la que el discurso tecnológico se enmarca. Para Chevallard, estos cuatro elementos en conjunto forman la noción de praxeología (que es, por definición, relativa a una institución) (Chevallard, 1990, [22]). Para esta teoría, los proyectos educativos se proponen la sucesiva organización de estructuras praxeológicas puntuales (aquellas orientadas a un tipo de tareas particular) en estructuras más complejas que permitan a les estudiantes operar sobre rangos amplios de situaciones problemáticas.
La idea de praxeología nos resulta interesante en la medida en que permite no sólo afinar la mirada sobre el trabajo de les estudiantes (identificando distintas dimensiones, algunas visibles y otras que pueden ser interpretaciones y que requerirán otros elementos para validarlas), sino poner de relieve que saber (Teoría-Tecnología) y saber-hacer (Técnica-Tarea) pueden construirse en conjunto, interrelacionadamente, y con referencia a determinadas cuestiones problemáticas.

Hasta aquí, hemos precisado cómo en nuestra concepción del aula como comunidad de producción, los problemas propuestos a les estudiantes cumplen un rol central en la construcción de conocimientos. Consideramos, siguiendo a otres autores como Sadovsky y Sessa (2005, [51]), que las interacciones que ocurren en el aula a propósito de la resolución de los problemas posibilitan una producción de conocimientos que solo pueden tener lugar en el espacio colectivo de trabajo.
Una parte de las interacciones, como hemos mencionado, serán del tipo estudiante-medio, a lo largo de la fase adidáctica, y en estas interacciones les estudiantes realizarán producciones (escritas u orales, completas o incompletas, matemáticamente correctas o erróneas) que indefectiblemente contendrán conocimientos que serán relevantes para el tratamiento en el aula. Brousseau (1976), citado por Artigue (2018) en este sentido, fue pionero al considerar que

> "El error no es simplemente el efecto de la ignorancia, de la incertidumbre, del azar, como

> lo creemos de acuerdo a las teorías empíricas o conductistas del aprendizaje, sino el efecto de un conocimiento anterior, que tenía su interés, su éxito, pero que ahora se revela falso o simplemente inadaptado."(P.6) [1]

Esta consideración nos sensibiliza sobre que el error tiene una dimensión de conocimiento, dotándolo de relevancia didáctica y matemática. Su carácter de inadaptado es la que nos brinda la oportunidad de investigar la relación del error con el problema en cuestión y, en particular, comprender en qué conocimientos se apoya le estudiante que lo produce y cómo se relacionan estos elementos con el conocimiento óptimo al que la situación didáctica apunta.

Ahora bien: consideremos un aula en la que les estudiantes trabajan en grupos de tres o cuatro personas a lo largo de una fase adidáctica. Cada grupo realizará producciones diferentes en cuanto a los conocimientos a los que apelan, a su validez matemática, al discurso utilizado, completitud de la respuesta[1], entre otras. Desde nuestra posición, es preciso que la enseñanza fomente la circulación colectiva de estas producciones, en instancias comunes gestionadas por le docente donde las mismas puedan ser objeto de discusiones, objeciones, ampliaciones y reformulaciones por parte de otres estudiantes y/o de le docente. Este espacio colectivo es también el ámbito para formular y contestar(se) nuevas preguntas que no surgen del trabajo autónomo de les estudiantes con el problema.
Partimos del supuesto de que fomentar estas interacciones genera buenas condiciones para que les estudiantes pongan en tensión su sistema de conocimientos, ya no solo al resolver un problema matemático, sino al considerar otras resoluciones que, como tales, apelarán a otros conocimientos, otras técnicas, otras maneras de comunicar, por mencionar algunos elementos. Estas instancias colectivas recibirán el nombre de “puesta en común” aunque disten considerablemente de una instancia de mera exposición de soluciones individuales.

Para poder cerrar esta sección, nos queda ocuparnos de un asunto central que dejamos abierto: supongamos que, en un aula en la que se ha propuesto una situación didáctica y a partir de la gestión de las producciones de les estudiantes se ha podido arribar colectivamente a la construcción de algunos conocimientos (aunque aún queden otros por construir y otros estén en proceso y requieran nuevas situaciones para ser retomados). Estos conocimientos, como ya mencionamos, estarán fuertemente contextualizados, y una marca de esa contextualización estará dada por el problema que dichos conocimientos permitieron resolver. Entonces, ¿qué acciones tomar, tendientes a que estos conocimientos adquieran un carácter descontextualizado (digamos, en algún sentido, más cercano a los saberes que se espera que les estudiantes hayan incorporado al finalizar el proceso de enseñanza-aprendizaje)?
La Teoría de Situaciones Didácticas (TSD) concibe en este sentido un tercer proceso, además del de devolución y de la fase adidáctica: el proceso de institucionalización, que incluye necesariamente la puesta

[1] Estas tres últimas características no son consideradas como absolutas sino que, por el contrario, son profundamente relativas a la comunidad de producción en cuestión, que en determinados momentos y en función de determinados consensos, tomará como correctas, adecuadas y completas a determinadas producciones que, en otro momento y/o en otra comunidad, no tendrán esa característica. Yackel y Cobb (1996, [59]) acuñan el concepto de “normas socio-matemáticas”, que consideramos muy potente para describir y estudiar estos fenómenos, pero que por motivos de extensión no desarrollaremos en este trabajo. Dentro de la TSD, un concepto crucial para mirar este asunto es la noción de Contrato Didáctico.

en común. En este, el rol del docente es de reorganizar, descontextualizar y enmarcar en el discurso de la disciplina los conocimientos y consensos construidos en el aula. En este proceso, que desde nuestra posición ocurre con plena participación de les estudiantes, tiene lugar la propuesta de definiciones de objetos matemáticos, la explicitación de sus propiedades y las relaciones que mantienen con otros objetos, así como la validación a través de demostraciones.

En resumidas cuentas: a la tarea docente de diseñar problemas y proponerlos en el aula, y a la tarea de gestionar las interacciones en el aula, se suma una tarea más: concebir acciones tendientes a descontextualizar el conocimiento producido y enmarcarlo en un discurso más amplio. Es en este discurso donde tiene lugar la construcción de la teoría matemática, en una versión cercana a la que se encuentra en el ámbito institucional.

En las siguientes secciones, pretendemos ahondar cómo estos asuntos se pondrán en juego en la propuesta didáctica que produjimos en el seno de esta tesis. Pero antes nos interesa hacer, aunque su brevedad no haga justicia a su valía y relevancia, tanto para la didáctica de la matemática en general como para esta tesis en particular, una mención a un último aporte teórico: la Ingeniería Didáctica (ID), cuya exponente principal es Michelle Artigue (1988, [3]).
La ID es, como tal, una metodología de investigación desarrollada en estrecha relación con la TSD. Artigue da una excelente caracterización de la intención primitiva de la que surge la ID:

> "[en Francia, a principios de los 80]Se denominó con este término a una forma de trabajo didáctico equiparable con el trabajo del ingeniero quien, para realizar un proyecto determinado, se basa en los conocimientos científicos de su dominio y acepta someterse a un control de tipo científico. Sin embargo, al mismo tiempo, se encuentra obligado a trabajar con objetos mucho más complejos que los objetos depurados de la ciencia y, por lo tanto, tiene que abordar prácticamente, con todos los medios disponibles, problemas de los que la ciencia no quiere o no puede hacerse cargo". (Artigue, 1988, P.2) [3]

La metodología que la ID propone es interesante, no solo en la medida que permite recortar y enfocar con un alto grado de detalle una determinada secuencia didáctica, sino que los constructos que propone para realizarlo permiten indagar en zonas del conocimiento matemático y didáctico que, entendemos, son relevantes como marco general para la elaboración de propuestas didácticas, sea o no que estas se ubiquen específicamente en este marco metodológico.
La metodología de la ID tiene cuatro fases que presentamos de manera simplificada:

- El análisis preliminar, que incluye un análisis "del campo de restricciones donde se va a situar la realización didáctica efectiva (...) distinguiendo tres dimensiones: la dimensión epistemológica (asociada a las características del saber en juego), la dimensión cognitiva (asociada a las características cognitivas del público al cual se dirige la enseñanza) y la dimensión didáctica (asociada a las características del funcionamiento del sistema de enseñanza)" (Artigue, 1988, P.4, [3])

- El análisis a priori, en el cual juega un rol particular el análisis de las posibles producciones de les estudiantes, de los conocimientos que estas pondrían en juego, y de cómo la gestión docente podrá gestionar y retomar estas cuestiones.

- La experimentación, donde la secuencia es llevada a cabo con la presencia de le investigadore, que recaba los datos.

- El análisis a posteriori, que permite la ponderación de los elementos del análisis a priori y de la experimentación para estudiar, discutir y eventualmente modificar la secuencia didáctica.

En el caso de esta tesis, que no se ubica en la ID en la medida en que no se produjo una implementación en aula -y, por lo tanto, no existe análisis a posteriori-. Sin embargo, consideramos que los elementos que hemos desarrollado en los capítulos 2 y 3, y los que desarrollaremos la siguiente sección, apuntan en la dirección de la primera fase, donde parte de la labor de le investigadore consiste en estudiar, para los conceptos que la secuencia se propone abordar, las significaciones y desarrollos de estos conceptos tanto histórica como actualmente. Recomendamos especialmente la lectura del artículo "Epistemología y Didáctica" ([1], donde la autora explicita su posición al respecto de estos asuntos.
En la tercera sección de este capítulo, presentaremos nuestra propuesta con un análisis que ubicamos en una dirección cercana a lo que propone la fase 2 de la ID.

4.2. Elementos específicos de la didáctica de la probabilidad y estadística

Un concepto que consideramos central para el abordaje de esta sección es el de modelización matemática (MM). Como refiere Patricia Sadovsky (2005):

> "Un proceso de modelización [matemática] supone en primer lugar recortar una cierta problemática frente a una realidad generalmente compleja en la que intervienen muchos más elementos de los que uno va a considerar, identificar un conjunto de variables sobre dicha problemática, producir relaciones pertinentes entre las variables tomadas en cuenta y transformar esas relaciones utilizando algún sistema teórico-matemático, con el objetivo de producir conocimientos nuevos sobre la problemática que se estudia." (P.27) [49]

Consideramos a la MM -concebida tanto para problemas en contexto intra como extra-matemáticos- como parte fundamental del trabajo matemático que nos interesa promover dentro del aula, independientemente del nivel educativo donde ésta tenga lugar. Por ello, este elemento atravesará la totalidad de la propuesta que diseñamos, entendiendo que es una práctica que se construye progresivamente a lo largo de una experiencia compartida de trabajo.

Destacamos especialmente, al igual que la autora, el carácter procesual de la MM. Es decir, no vemos a la MM como una actividad aislada que les alumnes realizan, para luego hacer otras tareas, sino que la entendemos como un proceso que posibilita -a la vez que condiciona y restringe- el trabajo matemático que les estudiantes llevarán adelante para resolver el problema. A su vez, este trabajo podrá redundar en que les estudiantes pongan en cuestión la pertinencia y la validez del modelo matemático propuesto.

Este posible diálogo entre modelo y problema nos parece particularmente fértil, especialmente en el seno de problemas de naturaleza estocástica.

Es central referir que la tarea de MM no está determinada completamente por el problema, sino que estará en estrecho vínculo con los conocimientos matemáticos de les estudiantes, y con las ideas (matemáticas o no) que elles tengan sobre el problema en cuestión. Esto lleva a considerar que estudiantes, problema y modelo matemático están implicados en una relación dinámica y compleja, cuyo desarrollo en el aula incidirá en qué conocimientos se construyan en el aula, y en los modos de esta construcción. Particularmente, aunque algunos problemas de MM pueden llevar a tareas que podrían abordarse descontextualizadamente en ciertos momentos, partimos del supuesto de que los sentidos construidos por los estudiantes en relación con estas tareas son diferentes a los que tendrían si se propusieran fuera del contexto de la MM.

Ahora bien, resulta pertinente preguntarse por la especificidad de la MM en el seno de problemas de naturaleza estocástica. Liliana Tauber (2001, [55]) refiere, en este sentido, citando a Willenski (1995, 1997):

> "En vez de encontrar la probabilidad a través de la resolución de conjuntos descontextualizados de problemas combinatorios, [les estudiantes] pueden participar en actividades constructivas, en las que diseñan y construyen modelos de probabilidad."(P.29) [55]

Pongamos, por ejemplo, un problema conocido como "La paradoja de Bertrand", propuesto por Joseph Louis Bertrand (1889, [8]), que consiste en preguntar por la probabilidad de que, dado un círculo y una cuerda de éste que sea elegida al azar, la longitud de ella sea mayor que la del lado del triángulo equilátero que puede ser inscripto en dicho círculo.

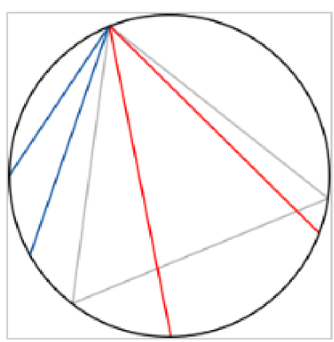


Figura 4.1: un triángulo equilátero inscripto en un círculo, y cuerdas de longitud mayor (rojas) y menor (azules) que el lado del triángulo.

Si lleváramos al aula este problema tal cual lo formuló Bertrand, cada estudiante o grupo de estudiantes podrá proponer algún modelo probabilístico para describir el fenómeno "elegir una cuerda al azar", potencialmente alguno de los siguientes:

- Elegir un punto uniformemente en el interior del círculo, y considerar la (única) cuerda cuyo punto medio es dicho punto. En este caso, la probabilidad del evento en cuestión es $\frac{1}{4}$

- Elegir un punto de manera uniforme en la circunferencia, y considerar la cuerda que lo une con uno de los vértices del triángulo. En este caso, la probabilidad del evento en cuestión es $\frac{1}{3}$
- Elegir uniformemente, en uno de los radios del círculo, un punto, y considerar la cuerda que es perpendicular a este radio. En este caso, la probabilidad del evento en cuestión es $\frac{1}{2}$

En un trabajo colectivo, le docente puede gestionar una discusión sobre la validez matemática de todos los modelos en sí mismos, mas no existen elementos en el enunciado (analíticos o experimentales) para decidirse por alguno de ellos. En parte, esto responde a que la propia formulación del problema es vaga con respecto a qué significa “elegir al azar”.
No obstante, cabe preguntarse por las características que puede adquirir este problema si puede ser planteado a partir de un trabajo de MM con relación a una situación experimental concreta. Para pensar en esta dirección, es interesante recuperar cómo, desde las ciencias físicas, se han dado numerosas discusiones a propósito de esta paradoja. Luego de un ida y vuelta que duró casi un siglo, Edwin. T. Jaynes (1973, [42]), logró dar con una solución relativamente cerrada al debate suscitado por la multiplicidad de estas soluciones, a partir de reformular el problema propuesto por Bertrand ligándolo a una situación experimental física: arrojar aleatoriamente un segmento de longitud L sobre un círculo de radio R, donde L es mucho más grande que R. Conviene señalar que esta propuesta le permitió reformular el problema y, con argumentos de naturaleza estrictamente matemática, determinar la probabilidad en cuestión con relación a ese experimento.
Consideramos que si les estudiantes producen un modelo probabilístico con referencia al experimento de tirar una varilla sobre un círculo, la repetición de los experimentos y su estudio estadístico podrá jugar un rol en la aceptación o no del modelo. A la vez, el estudio de la experimentación puede generar nuevos conocimientos sobre el modelo, como por ejemplo, la posibilidad de que en algunos casos la varilla no corte el círculo en dos lugares. Esta posibilidad requeriría una reformulación del modelo inicial.

Desde nuestro punto de vista, vemos importante la incorporación de experimentaciones que permitan generar datos, ya que su estudio puede ser una apoyatura para el establecimiento de nuevas relaciones sobre la problemática estudiada, y eventualmente la modificación del modelo probabilístico. En este sentido, como refiere Fischbein (1975, [33]),

> “La experiencia estocástica es a la probabilidad lo que la experiencia espacial es a la geometría. En ambos casos, aunque finalmente se llegue a una forma estrictamente axiomática de la exposición, los fundamentos, los problemas, la naturaleza de las soluciones y el estilo de pensamiento están estrechamente ligados al carácter específico del dominio. La construcción del concepto de probabilidad parte de experiencias concretas de carácter estocástico” (P. 16) [33]

Desde el punto de vista del diseño de esta propuesta didáctica y, más en general, desde una perspectiva didáctico-matemática, nos interesa sostener y promover la articulación, para una problemática dada, entre un modelo probabilístico y un conjunto de experimentos repetibles asociados a la misma. Concebimos una bidireccionalidad en esta articulación: por un lado, el modelo probabilístico podrá funcionar como soporte para la confección y evaluación de la pertinencia de los experimentos propuestos; por otro lado, el análisis estadístico de los resultados proporcionará información para operar sobre el modelo mediante reajustes o reformulaciones. El trabajo matemático articulado entre el modelo y los experimentos podrá funcionar como sostén para la elaboración de respuestas en torno a la problemática. A fines de nuestro trabajo, estos procesos particulares de MM recibirán el nombre de Procesos de Modelización Estocástica, o Modelización Estocástica[2] a secas, sobreentendiendo su carácter procesual.

Conviene señalar que, dado un problema susceptible de admitir una Modelización Estocástica, la relación entre el modelo probabilístico, la experimentación y el problema podrá darse de maneras diversas en el aula. Tomemos por caso el problema 1 de nuestra propuesta (ver sección siguiente), en el que la tarea para les estudiantes es predecir el comportamiento de un dado que (sin que elles lo sepan de antemano) está cargado. Antes de proponer un modelo probabilístico, les estudiantes pueden realizar el experimento de lanzar el dado repetidamente, y proponer un modelo probabilístico -apoyado en el análisis de los resultados- en el que se formule explícitamente la carga de cada una de las caras. Este es un ejemplo en el que el la modelización estocástica puede configurarse progresivamente, originalmente apoyada en la modelización de una problemática para la cual la experimentación ya está dada.

Por otro lado, la complejidad técnica de algunos experimentos o la cantidad de los mismos que se desee hacer puede requerir o hacer preferible el uso de simulaciones. Como refiere Luis Santaló (1980)

> "tanto en los ejercicios hechos en clase, como en los proyectos fuera de ella, las experiencias reales pueden ser engorrosas y difíciles de realizar, siendo preferible la simulación por experiencias ideales." (P.205) [50]

Nos interesa traer la posición de Carmen Batanero (2005, [4]), que pone matices a la relación experimento-modelo, planteando que:

> Un enfoque experimental puro de la enseñanza de la probabilidad no es suficiente incluso cuando la simulación nos ayuda a encontrar la solución a problemas de probabilidad que surgen de la vida real (...) porque la simulación depende de las hipótesis establecidas de antemano y del modelo teórico que implementamos en el ordenador. Un conocimiento genuino de probabilidad solo se alcanza con el estudio de alguna probabilidad formal, aunque debe ser gradual y estar apoyado en la experiencia estocástica de los estudiantes." (P. 260) [4]

Todos estos elementos teóricos nos plantean el desafío de diseñar problemas para el aula de matemática que movilicen la producción de Modelos Estocásticos y fomenten -en clave de lo desarrollado en la sección anterior- la construcción de conocimientos orientados al desarrollo de teoría probabilística. En el Capítulo 2 esbozamos la relación histórica entre el desarrollo de conceptos probabilísticos y la solución de problemas provenientes de la inferencia estadística. Esto nos permitió preguntarnos por la

[2]Algunos trabajos muy recientes, como por ejemplo Esteley y Magallanes (2021a y b, [31], [32]) y , se refieren a la Modelización Estocástica en un sentido cercano al que tomamos en este trabajo.

posibilidad, a la hora de diseñar los problemas, de apelar a la inferencia estadística como fuente para la propuesta de problemas en el aula. En esta dirección también avanzan Batanero y Borocvnik (2016, [5]):

> "La estadística descriptiva, la teoría de la probabilidad y la inferencia estadística se complementan y ven la misma información desde ángulos diferentes. La estadística descriptiva investiga la información de una muestra (conjunto de datos) y la resume mediante números únicos y representaciones gráficas. La inferencia estadística va más allá del conjunto de datos actual e intenta generalizar los resultados, es decir, transferirlos a una población más amplia a la que pertenece el conjunto de datos. La probabilidad desempeña el papel de mediador, ya que proporciona una justificación para una generalización más allá de los datos iniciales" (P. 2 y 3) [5]

Esta interrelación entre la Teoría de Probabilidades, la Inferencia Estadística y la Estadística Descriptiva (de ahora en más, la terna TP-IE-ED), se nos muestra fértil y potente para el diseño de problemas. De hecho, todos los problemas de nuestra propuesta didáctica abordan problemas ubicables en el terreno de la inferencia estadística.

Ahora bien: el movimiento que lleva del trabajo colectivo a propósito de problemas matemáticos a la construcción de una TP apelando al posicionamiento que venimos proponiendo supone, para la enseñanza, la exposición a una serie de tensiones que no son elementales de resolver.

A fines ilustrativos, quisiéramos mencionar cómo proponemos abordar la Ley Débil de los Grandes Números en el contexto de esta propuesta. El problema 4 de esta propuesta propone, inspirado en la situación fundamental propuesta[3] por Brousseau et al (2002, [15]), a les estudiantes la tarea de estimar la composición de una urna con 100 bolillas, de las cuales algunas son blancas y otras negras. Les estudiantes podrán producir estimaciones numéricas a partir de repetir el experimento "extraer una bolilla, anotar si es blanca o negra, y luego devolverla" y luego promediar la cantidad de veces que obtuvieron una blanca. Estas estimaciones, para cantidades grandes de repeticiones, involucrarán números con coma, los cuales les estudiantes redondearán a fin de proponer un valor estimado para la cantidad de blancas de la urna. En función de este hecho, en el problema se pide que les estudiantes estimen la probabilidad de que, entre la proporción real y la estimación, la diferencia no sea mayor que $0,005$[4].

Una vez enfrentades a esta situación, excepto para valores de N relativamente pequeños, el conocimiento óptimo al que deberán apelar para estimar la probabilidad de acertar es la desigualdad de Bienaymé-Tchebychev. Hacia el final del problema 4, proponemos que el abordaje colectivo de la demostración de la Ley Débil de los Grandes Números pueda apoyarse en esa experiencia de trabajo.

[3]Para estudiantes de nivel primario, a propósito de la noción de probabilidad.

[4]Cabe mencionar que, si bien la posibilidad de "acertar"no forma parte de la cultura estadística tradicional, en el contexto muy específico de este problema este evento puede ser interpretado como la probabilidad de acertar. Omitimos formularlo en estos términos para evitar proponer para el aula prácticas que no se retomarán en el ámbito de la estadística, a pesar de que el modelo propuesto junto con la manipulación analítica generada por los redondeos, dotan de sentido a este evento en el marco donde se desarrolla

Desafortunadamente, en el desarrollo de esta tesis nos encontramos frente a la dificultad para concebir uno o más problemas que apunten a que les estudiantes conjeturen o validen la desigualdad antes mencionada. Asimismo, dada la complejidad del problema 4, no encontramos pertinente ninguna posible adaptación que permita también conjeturar y/o validar esta desigualdad. No tomamos posición respecto de si esta desigualdad debería quedar a cargo del discurso de la enseñanza, o habría que seguir buscando un “buen” problema para abordarla. Algo similar ocurre con el Teorema Central del Límite. En el problema 5 de esta propuesta, con menos ambición que en el problema 4, nos propusimos que les estudiantes puedan arribar a una formulación del teorema, a la construcción de sentidos en torno al factor “raíz de N” con una primera experiencia en este problema, a partir del estudio del caso en el que se suman variables normales independientes idéndicamente distribuídas, y del cálculo, en el caso i.i.d, de la varianza del promedio de variables aleatorias presente en la estandarización, y a la idea de que el promedio estandarizado “cada vez se parece más” a una normal. No obstante y como desarrollamos en el capítulo anterior, las diferentes demostraciones de este teorema involucran técnicas específicas que son particularmente complejas de referenciar o construir en el seno de trabajos de Modelización Estocástica, por lo que anticipamos que la demostración de ese teorema quedará en buena medida del lado de un discurso organizado mayoritariamente por le docente. La simulación computacional es una herramienta que permite explorar el comportamiento de muchos fenómenos, generando conjeturas que pueden ser luego demostradas. Si bien la simulación no necesariamente ilumina las razones matemáticas que que permiten demostrar resultados, juegan un papel fundamental para que les estudiantes puedan apropiearse de los objetos involucrados. En particular, concebir, manipular y entender la distribución del promedio es un desafío para la disciplina, más allá de poder operar con el de manera analítica.

Estas consideraciones nos llevan a querer resaltar la necesaria articulación entre la resolución de problemas y el desarrollo de la TP, de modo que se maximicen las oportunidades de que los constructos específicos de la TP cobren sentido para les estudiantes en relación con dichos problemas, y que estos puedan motivar los desarrollos, la formulación de conjeturas, la utilización de técnicas empleadas en las demostraciones, entre otros elementos que aparezcan durante el desarrollo de la TP.

4.3. La propuesta didáctica

En el primer apartado de este capítulo nos hemos referido a una trama matemático-didáctica que involucra las interacciones del aula, los conocimientos de les estudiantes, la red de saberes institucionales, el rol de le docente, entre otros elementos. También hemos profundizado en torno a la especificidad de cuestiones para el abordaje de asuntos concernientes a la estocástica. En particular, hemos mencionado que partimos de la premisa constructivista, según la cual les estudiantes aprenden en interacción con un medio que ofrece resistencias, y que este medio se conforma a partir de una problemática a resolver y unos conocimientos que evolucionan a lo largo de la situación didáctica y la interacción los otros actores de la clase.

Antes de abocarnos a un análisis en fino de nuestra propuesta didáctica, queremos dar robustez al

sentido que “propuesta didáctica” cobrará en el contexto de esta tesis de licenciatura, y en particular de esta última sección.
Este concepto, desde nuestro punto de vista, tiene tres partes:

1. Un conjunto secuenciado de problemas, tareas o preguntas para proponer en un aula.
2. Una intencionalidad didáctica.
3. Un análisis didáctico-matemático de los problemas y tareas que permita anticipar resoluciones de les estudiantes y gestiones de le docente, estableciendo un vínculo entre los problemas y la intencionalidad didáctica.

Ante todo, nuestra idea de propuesta asume una posición no prescriptiva: no pretende erigirse como mejor que otras, ni como la única para perseguir los fines propuestos, ni ninguna otra cuestión que le adjudique, en algún sentido, una preeminencia sobre otras. Sí apostamos a que la manera en que fue construida y los elementos puestos en juego inviten a discutir con ella, a probarla, a mejorarla.

En términos de 1., nuestra propuesta está conformada por cinco problemas con unas determinadas acciones y/o intenciones docentes tanto para su puesta en aula, como para los espacios de trabajo que se abren entre el fin de uno y el comienzo del siguiente.

En términos de 2., nuestro objetivo es que les estudiantes -en principio de las materias de probabilidad y estadística de nuestra facultad- construyan conocimientos estocásticos en torno a cinco grandes asuntos matemáticos:

- El concepto de probabilidad, tanto frecuentista como axiomática.
- La noción de independencia.
- La noción de variable aleatoria.
- Ley Débil de los Grandes Números.
- El Teorema Central del Límite.

Para la realización del análisis propuesto en 3., nos apoyaremos en las producciones realizadas en esta tesis, a saber: el análisis histórico-epistemológico esbozado en el capítulo 2, la discusión propuesta en el capítulo 3, y los elementos esbozados en las secciones anteriores de este capítulo.

En las siguientes líneas, propondremos los enunciados de los problemas y realizaremos un análisis didáctico-matemático de los mismos. También explicitaremos lineamientos generales de los espacios de trabajo que consideramos que deberían existir entre el fin de un problema y el inicio del siguiente. Excede a la extensión y dimensión del trabajo de esta tesis diseñar una propuesta que logre abarcar el completo de asuntos que consideramos que quedan comprendidos entre el problema 1 y el problema 5.

No obstante, tenemos la esperanza de que los análisis didáctico-matemáticos propuestos, los lineamientos que esbozaremos y los elementos que desplegamos en esta tesis puedan funcionar como un soporte sólido para que, cada docente que quiera llevar a sus aulas versiones personales de esta secuencia, pueda hacerlo.

En torno al análisis didáctico-matemático de cada problema, este contendrá esencialmente cuatro ejes interrelacionados:

- Conocimientos que les estudiantes podrían tener disponibles para abordar el problema.
- Anticipaciones de posibles resoluciones (erróneas, o no) que les estudiantes podrían desplegar al intentar resolver el problema.
- Posibilidades para la gestión docente de las interacciones en el aula, tanto en pequeños grupos como en las instancias de discusión colectiva.
- Conocimientos nuevos que puedan construirse a lo largo de la situación didácitca, incluyendo tanto las producciones matemátics de les estudiantes como las interacciones en el aula a propósito de estas producciones.

Los problemas propuestos están diseñados con la intención de que sean abordados tanto en instancias de trabajo autónomo en pequeños grupos, como en instancias colectivas a propósito de las producciones de los grupos. No ubicamos, como referimos más arriba, a la instancia colectiva como una instancia de mera reunión de soluciones al final del trabajo autónomo. En todos los problemas propuestos se prevén puestas en común, que jugarán un rol importante en el despliegue del problema en el aula. Las consideramos como oportunidades para que ideas matemáticas (algunas nuevas, otras no, algunas correctas, otras no) que le docente anticipará (¡o no!) se sometan a un tratamiento colectivo que se concibe como parte del trabajo con el problema. También, como oportunidades para que le docente incorpore elementos, escrituras, argumentos que les estudiantes no hayan desplegado pero que sean fértiles tanto para complejizar las producciones de elles, como para abordar los ítems siguientes. En particular, en el ámbito de problemas de Modelización Estocástica, en estas instancias pueden ponerse en común los resultados (esperablemente diferentes) de las experimentaciones de cada grupo, lo que según el problema posibilitará discutir y abordar diferentes asuntos.

Ahora bien: la (necesaria) interacción de le docente con los conocimientos matemáticos de sus estudiantes no podrá ocurrir de otro modo que no sea a partir de las producciones matemáticas de elles (ya sea que estas sean accesibles a le docente de manera escrita u oral). En este sentido, el análisis -anterior a la puesta en aula- de posibles resoluciones cumple un lugar central al darle a le docente una gama de anticipaciones que puedan ampliar su margen de acción para interpretar las producciones de les estudiantes. Nuestro análisis también incluye, además de maneras correctas y erróneas, completas e incompletas de resolver el problema, una mirada didáctico-matemática sobre los conocimientos que pueden subyacer a ellas, y una posición respecto de cómo abordarlos tanto individual como colectivamente. A su vez, este trabajo anticipatorio cumple otro rol fundamental: el del análisis del -posible-funcionamiento del propio problema en el aula, lo que permite someterlo a reconfiguraciones y ajustes en su enunciado y gestión prevista, de acuerdo a los datos contextuales en los que el curso sería llevado a cabo.

Sostenemos la potencialidad de la propuesta de trabajo en pequeños grupos a propósito de los problemas. Entendemos que la producción de un pequeño grupo puede (y debe) redundar en que les estudiantes discutan y argumenten matemáticamente de manera autónoma. Pensar colectivamente un problema puede implicar poner al servicio de una solución nuestras ideas, concepciones e intenciones, y someterlas asimismo al juicio de pares que comparten el proyecto de solución, movilizando y tensionando nuestro sistema de conocimientos en esta interacción. Estos elementos posiblemente se sinteticen en una única solución que el grupo compartirá en la instancia de puesta en común, y quedará para la gestión docente la tarea (para nada trivial, y que esperamos pueda apoyarse en el análisis desplegado) de articular estas producciones en el seno del trabajo en el gran grupo.
El trabajo matemático que proponemos y promovemos (que esperamos haya empezado a delinearse en las líneas anteriores) supone una dinámica de circulación del conocimiento en el aula que no necesariamente es habitual, y que por ende requerirá una construcción colectiva que a su vez es progresiva. Que para un determinado problema sean admisibles respuestas diversas (y, en particular, que sean objeto de reflexión y análisis tanto aquellas correctas como incorrectas), que en la lógica de trabajo no necesariamente exista un contenido previo desarrollado específicamente para el abordaje de los problemas, y que se sostengan discusiones matemáticas entre estudiantes y docentes como fuente sustantiva de conocimiento y relaciones matemáticas, son prácticas muchas veces nuevas y que requieren una construcción también dentro del aula.

Habiendo desplegado estos elementos, nos interesa abordar el análisis didáctico-matemático de la propuesta, en particular de los problemas que la componen.

4.4. Los problemas de la propuesta

4.4.1. Problema 1

Este primer problema se concibe como punto de partida para abordar la noción de probabilidad, y se considera posible como primera instancia de trabajo en el curso. Es parte de la intención que les estudiantes aborden los ítems a partir de sus ideas en torno a la probabilidad y el azar, sin necesariamente tener una definición rigurosa de ninguna de ellas.
En una primera etapa del problema, apuntamos a llegar a una noción de probabilidad asociada a la factibilidad de ocurrencia de un fenómeno, donde la idea de factibilidad está cuantificada por la proporción esperada de veces que, en una gran cantidad de repeticiones, ocurrirá ese fenómeno. A partir de este trabajo en el aula, se espera que les estudiantes puedan apoyarse en la probabilidad estimada de un evento, generar anticipaciones respecto de lo que ocurrirá en una gran cantidad de repeticiones del experimento correspondiente. También, en este problema se apunta a discutir la relación entre la canti-

dad de repeticiones del experimento y los resultados en cada una de ellas, relacionando la discrepancia entre las anticipaciones y la experimentación con la idea de variabilidad de una muestra. Finalmente, en este problema se avanza en la utilización de propiedades de aditividad de la probabilidad y en el cálculo de la probabilidad de eventos de cardinal mayor a 1 a partir del cálculo de probabilidades de eventos simples.
A cada grupo de tres o cuatro estudiantes, le docente da un dado desequilibrado (los números 2 y 5 tienen probabilidad de ocurrencia $p = \frac{1}{4}$, y el resto $\frac{1}{8}$). Ante la falta de disponibilidad del dado físico, se puede dar un simulador [5]. Se otorga también una hoja donde están puestos los resultados de 24 tiradas del dado (diferentes para cada grupo). No se les informa a les estudiantes que el dado está o no cargado.

Enunciado y Análisis del Problema 1

En un casino de Buenos Aires, su grupo va a jugar un juego de azar. Este juego consiste en anticipar los resultados que dará un dado cuando un croupier lo tira sobre la mesa. Los resultados de las primeras 24 tiradas son los que ustedes tienen en la hoja.
a) Elijan los dos números que creen que saldrán más veces en las siguientes 48 tiradas. Escriban por qué los eligen.
Importante: Todos los grupos del aula juegan con el mismo croupier, solo que en diferentes ocasiones.

Conviene señalar que, en línea con lo que referimos líneas más arriba, será muy importante el proceso de devolución de esta consigna hasta lograr un consenso respecto de la situación que está presentando el problema (en particular, podría ser necesario aclarar qué es un croupier).
En caso de que le docente dé los dados, debe quedar claro que son una copia exacta del dado del croupier, y que las 24 tiradas fueron realizadas en momentos diferentes para cada grupo. En caso de dar el simulador, deberá aclarar que el mismo simula tiradas del dado del croupier, y que todos los grupos tienen el mismo simulador. Y pedirles que se abstengan de utilizarlo hasta el ítem b). Teniendo en cuenta que este problema está pensado como para comenzar el trabajo en un curso de probabilidad y estadística, es posible que los conocimientos que les estudiantes tengan sobre problemas de azar esté directamente referido a sus propias experiencias con los juegos de azar, y no a una teoría matemática desarrollada para modelizarlos. Sería objeto de un estudio en sí mismo ponderar de manera más o menos exhaustiva los diferentes estados de conocimiento con que les estudiantes podrían abordar estos problemas, pero asumimos que elles conocerán algún juego que implique arrojar dados (por ejemplo, la generala).
En este sentido, las ideas sobre juegos con dados seguramente estén apoyadas en dichas experiencias, en las que los dados seguramente no están cargados, y por tanto puede ser que arriben al problema con esta idea. Esto podría llevarles a pensar que los 24 datos son irrelevantes, y que pueden elegir cualesquiera dos números para su respuesta[6].

[5]Uno posible es el que se encuentra en el siguiente enlace, eligiendo la opción “2-5 flat” y seleccionando n=1.: https://www.randomservices.org/random/apps/DiceSample.html.

[6]No obstante, es pertinente observar que, si les estudiantes cuentan con el dado físico, es posible que al percibir su carga se pregunten por las características del dado, que será “diferente” a los dados con los que ya han experimentado.

Anticipamos que una gran cantidad de estudiantes va a considerar relevantes los datos ofrecidos, y lo tendrán en cuenta en su respuesta, aunque posiblemente de diferente modo.
Por ejemplo: algunes de elles, presuponiendo el dado equilibrado, podrían considerar que, si en los datos figura muchas veces número 5, es las siguientes 48 tiradas deberá salir menos veces. Esta idea es un poco cercana a la que se conoce en la bibliografía como falacia del jugador.. Estes estudiantes elegirían los dos números que menos salieron.

Con soporte en los mismos datos, pero razonando diferente, podría ser que les estudiantes consideren que, si bien lo que pasó antes no condiciona lo que pasará después, sí informa sobre cómo es el comportamiento del dado.

Entonces, una aparición de muchos números 5 puede indicar una mayor propensión a que aparezcan nuevamente, y entonces elijan al 5 como uno de ellos, y al otro número como aquel que aparezca más veces, después del 5. Esta posición podría habilitar la suposición de que el dado está desequilibrado. Finalmente, debemos considerar posiciones mixtas posibles, como ser: considerar, para la elección, el número que más haya salido y el número que menos haya salido, siguiendo lógicas diferentes en la misma elección.

Run	X1	Run	X1
1	6	25	5
2	2	26	1
3	1	27	3
4	2	28	1
5	3	29	2
6	3	30	5
7	1	31	5
8	6	32	3
9	5	33	4
10	2	34	3
11	1	35	4
12	5	36	4
13	5	37	1
14	2	38	2
15	5	39	4
16	6	40	3
17	5	41	3
18	5	42	3
19	4	43	5
20	3	44	2
21	5	45	5
22	5	46	1
23	2	47	3
24	6	48	1

Figura 4.2: Dos tiradas de 24 dados

Un elemento particular de este problema está dado por la poca cantidad de repeticiones, lo que hace que la variabilidad de la estimación sea alta, lo que redunda en que utilizando mismos criterios, arriben a elecciones diferentes.
En la figura 5.2 se ven dos tandas de 24 tiradas. En la primera tirada el 2 y el 5 son los valores más repetidos, aunque el 6 es el tercero por apenas una vez. En la segunda tanda hay apenas 8 ocurrencias del 2 o del 5, y los valores más repetidos son el 1 y el 3. Vemos cómo la baja cantidad de repeticiones incide en la variabilidad de las estimaciones.

Consideramos que puede ser interesante que, luego de realizar este ítem, le docente pueda gestionar una instancia en común para que cada grupo comparta sus respuestas, el criterio que utilizaron para decidir y por qué confían en ese criterio. El hecho de que todos los grupos jueguen con el mismo croupier y "con el mismo dado", podría generar extrañeza en el caso de que, por ejemplo, con el mismo criterio dos grupos diferentes hayan elegido pares diferentes de números. También podría pasar que esto pase desapercibido, aunque en este caso consideramos pertinente que le docente instale esta cuestión. Les estudiantes que asuman equiprobabilidad, probablemente no encuentren conflictiva esa situación, por apelando auna idea como ser "jugaron en ocasiones diferentes, entonces los resultados serán diferentes". En la discusión sobre posibles motivos, podría surgir como una posible que el dado

esté cargado. Esto podría generar algo de revuelo en el aula, especialmente para quienes asuman equiprobabilidad. Le docente, consideramos, debe abstenerse de aportar información en este sentido. En el pizarrón, para cada grupo, le docente podrá anotar su elección y una explicación del criterio utilizado para esta elección. Es posible que la diversidad de criterios y la variabilidad de las muestras presente una disparidad tal que no pueda arribarse a una respuesta consensuada grupalmente, lo cual nos parece rico ya que permite que cada grupo pueda sostener sus elecciones y someterlas a evaluación a lo largo de los siguientes items. Anticipamos que podría pasar que, a partir de estas discusiones, algunos grupos de estudiantes cambien de estrategia de cara al ítem b).
En esta puesta en común el objetivo no es decidir si un criterio está bien o mal. Puede ocurrir que haya mayor o menor consenso respecto de algunos criterios, consenso que puede modificarse con referencia a los resultados de las siguientes experimentaciones. La intención es que, con el correr de los siguientes dos ítems, pueda arribarse a que la mejor estrategia es elegir los dos números que más hayan salido, aunque esta estrategia no garantiza que la elección y los resultados siguientes se condigan. También, que cuantos más resultados se tengan en cuenta, mayor confianza habrá en esta estrategia. El conocimiento matemático que subyace este criterio es la Ley Débil de los Grandes Números.

b) Realicen las 48 tiradas. ¿Acertaron en su elección? Ahora, conociendo 72 tiradas del dado, ¿cuál sería su elección para las próximas 48 tiradas?

De acuerdo con la variabilidad y diversidad que referimos anteriormente, es esperable que la mayoría haya perdido y que los grupos que acertaron sean aquellos que hayan elegido el 2 y el 5, independientemente de si lo hicieron con el criterio de mayor frecuencia o no.

La segunda parte del ítem apunta a que, al tener 72 repeticiones del experimento y analizar sus resultados, el criterio que utilizaron para el primer ítem determine una elección diferente que la ya realizada. Para les estudiantes que aún asuman la equiprobabilidad, posiblemente sostengan que perdieron "por mala suerte" o que ganaron "por buena suerte", y entonces su apuesta nuevamente sería construida con el mismo criterio. Vemos fértil que le docente les proponga que estudien los valores de las 72 jugadas para evaluar si todos los números salieron una cantidad parecida de veces, con la intención de fomentar que les estudiantes relacionen sus supuestos sobre el comportamiento del dado con la información que la experimentación provee.

La confianza que tengan les estudiantes en su criterio podrá determinar si lo sostienen para la nueva decisión. De hecho, es improbable, pero no imposible, que en un grupo un criterio desacertado utilizado en el ítem a) arroje como resultado una elección correcta en función de lo experimentado en el ítem b), por lo cual posiblemente sostendrían este criterio para la siguiente tirada (o la elección).

En este contexto, una puesta en común puede ser interesante para volcar en el pizarrón las nuevas situaciones: qué grupos ganaron o perdieron, y qué decidieron en función de dicho resultado y de los datos (incluyendo tanto la nueva elección de números, como el criterio utilizado). Anticipamos que la

gran mayoría habrá perdido, y que esto producirá algún tipo de sorpresa o malestar. Y que posiblemente los grupos que hayan ganado sean les que eligieron el par (2,5) en el ítem a). En este caso, sería interesante que le docente pregunte al resto de los grupos qué hubiera pasado si elegían estos dos números (anticipamos que la mayoría habría ganado en este caso). Otro asunto importante para discutir en esta puesta en común, además de "cuál era el par correcto para elegir", qué ideas matemáticas subyacen a cada criterio. Por ejemplo, el criterio que apela al descarte de los números que más salieron, puede ya discutirse en términos de los supuestos que la subyacen.
Con esta puesta en común, consideramos que se arribará a un consenso general de que el 2 y el 5 parecen tener una mayor preponderancia a salir en el dado del croupier. Nos parece importante que le docente pueda discutir con les estudiantes que es esto de la "mayor preponderancia". Seguramente algunes estudiantes apelen a que "estos dos números salen más veces que el resto". Le docente tendrá la intención de cuantificar la relación entre la cantidad total de tiradas, y la cantidad total de veces. La elección de la probabilidad para los números 2 y 5 permite que, para valores como 72, la cantidad de veces que salen sea aproximadamente la mitad. Con esto en mente, le docente puede proponer el item c).

c) Completen la siguiente tabla con los datos de su dado (para las columnas con valores más grandes, utilicen un simulador):

Cantidad de tiros	**24**	**48**	**72**	**500**	**1000**
Porcentaje de veces que sale el 2 o el 5					

Aquí, cada grupo deberá completar la tabla con sus datos experimentales. Esta tarea no tiene una gran cantidad de maneras de ser resuelta, aunque sí es interesante que le docente tome los resultados de cada grupo y los ponga en común. En todos los casos, si bien los primeros valores de la columna serán bien diferentes, los correspondientes a la última serán parecidos entre sí. Le docente podrá preguntar entonces por la equiprobabilidad o no de un dado. Si bien consideramos que para este momento la mayoría de les estudiantes estará convencide de que el mismo está cargado, una buena pregunta que le docente podrá instalar es "si el dado no estuviera cargado, ¿qué resultados deberíamos ver en la última columna de la tabla? Esta pregunta podrá ser respondida mediante un simulador, o con un razonamiento de orden más probabilístico: si ninguno tiene preferencia por salir, más o menos saldrán todos la misma cantidad de veces, y luego entre el 2 y el 5 habrá aproximadamente la tercera parte de 1000, o sea 333, en vez de aproximadamente 500, como en este caso. Esta discusión puede ser un buen insumo para que les estudiantes trabajen autónomamente en torno al último item:

d) Supongamos que tiramos 2000 veces el dado. Anticipen cuál o cuáles números aparecerán menos veces. ¿Cuántas aproximadamente?

En este último ítem, cada grupo de estudiantes deberá mirar el comportamiento del dado ahora en números que no son ni el 2 ni el 5. Será interesante que el trabajo autónomo de los grupos tenga un buen espacio dentro del aula, a fin de que les estudiantes puedan poner en relación lo trabajado hasta el momento, sus ideas y la pregunta del item d).

El estudio de los valores 1, 3, 4 y 6 para las 1000 tiradas del item anterior podrá arrojar diferentes interpretaciones. Algunes estudiantes podrían razonar que, como la mitad de las veces sale el 2 o el 5, la otra mitad de veces sale cualquiera de los otros números, involucrando una asunción de equiprobabilidad para los otros números restantes. La apoyatura de este argumento en la experimentación será reinterpretando los datos experimentales ya analizados.
Otra estudiantes podrían apelar a un estudio exhaustivo de cada uno de los otros números, posiblemente anticipando que, así como el 2 y el 5 salen más que el resto, podría pasar que alguno de los otros 4 salga más que los otros. Según cada muestra simulada en el ítem anterior, habrá uno, o a lo sumo dos de estos cuatro valores que habrán salido menos veces. Aunque las cantidades que sale cada uno de ellos serán similares entre sí. Algunes estudiantes se apoyarán en la contingencia de que, por ejemplo, el 3 haya salido 123 veces, mientras que el 1 126, para afirmar que el 3 es el menos probable. Otres quizás sostengan la hipótesis de que cualquiera de los cuatro números sale con igual preponderancia, ya que la diferencia entre las veces que salen unos y otros es muy pequeña si se la compara con la cantidad total de tiros.
Es interesante señalar que estas tres ideas, con diferente grado de correctitud matemática, permitirán que les estudiantes, apoyades en lo realizado en el ítem c), concluyan correctamente la cantidad de veces esperada para el o los valores menos probables, ya sea dividiendo a la cuenta 1000/4 (representando la repartición equitativa de los aproximadamente 1000 valores que no serán ni el 2 ni el 5), o reescalando la relación obtenida experimentalmente (123 entre 1000) a 2000 tiradas, ya sea con una regla de tres simple, o duplicando el 123, o calculando 123/1000, y luego multiplicando por 2000.

Le docente podrá, en la puesta en común, comenzar preguntando solamente las estimaciones que cada grupo dio, y para cuál/cuáles valores, para luego indagar más exhaustivamente en los argumentos que hayan sustentado dichas aproximaciones. Nos interesa que le docente pueda recuperar los argumentos y someterlos a tratamiento colectivo, por ejemplo comenzando por el argumento que asuma la equiprobabilidad de los cuatro valores restantes (anticipamos que les estudiantes no hablarán en términos de equiprobabilidad, sino de que cualquiera puede salir con la misma preponderancia). Si bien la conclusión es correcta, la premisa podría ser objetada por el resto de les estudiantes o por le docente, que podría apelar ahora a incorporar el razonamiento de alguno de los grupos que hayan mirado exhaustivamente los otros cuatro valores. Lo interesante de esta puesta en relación será que esta mirada exhaustiva permitirá sustentar la posición del primer grupo. Le docente podrá comenzar, a la vez que finaliza la instancia de trabajo sobre este problema, dar lugar a una primera instancia de institucionalización. Un punto de partida posible para ello es escribir en el pizarrón las proporciones en las que cada número salió, con una escritura como la siguiente:

$$\text{Sale 1} \;\text{——}\; \frac{1}{8}\,;\text{Sale 2 o 5} \;\text{——}\; \frac{1}{2}\,;\ \text{Sale 3} \;\text{——}\; \frac{1}{8}\,;\ \text{Sale 4} \;\text{——}\; \frac{1}{8}\,;\ \text{Sale 6} \;\text{——}\; \frac{1}{8}$$

y ya introducir en el aula la noción de probabilidad de un evento, asociada a la proporción de veces que veces que, para una gran cantidad de repeticiones del experimento, se espera que este ocurra. Esta idea de “gran cantidad de repeticiones” puede apoyarse también en una escritura tipo límite. La idea de

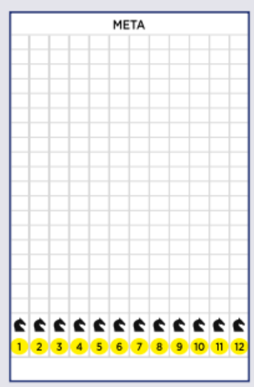
evento podrá definirse como un subconjunto de Omega, definido como el espacio de todos los resultados posibles del experimento.
Observemos que le docente podrá preguntar por cómo calcular la probabilidad de que salga el 5, anticipando que algunes estudiantes digan que es $\frac{1}{4}$, por dividir al $\frac{1}{2}$ entre dos, o bien volviendo a los datos experimentales. Nuevamente, el primer razonamiento asumirá una equiprobabilidad entre el 2 y el 5, que solo podrá sostenerse en los datos experimentales. Será interesante que le docente pregunte entonces qué relación hay entre la probabilidad de que salga 2, la probabilidad de que salga 5 y la probabilidad de que salga uno de los dos, pudiendo arribar que es la suma de las probabilidades de ambos. Esta propiedad podría ser validada a partir de la definición de probabilidad ya esbozada, y configurar en esta instancia de institucionalización en la formulación de la propiedad de la probabilidad de la unión disjunta.

▲

Al finalizar este problema, con su instancia respectiva de institucionalización, habrán quedado explicitas: una primera definición frecuentista de probabilidad, una técnica de estimación de la probabilidad a partir del estudio de frecuencias, y una técnica de cálculo de probabilidad de eventos disjuntos como suma de probabilidades de eventos unitarios. Consideramos que en función de lo trabajado, podrían abordar un siguiente problema en un escenario que involucre a un dado equiprobable y al cálculo de probabilidades de eventos asociados, con la intención de avanzar en la discusión de técnicas para el cálculo de probabilidades de eventos en el contexto de espacios equiprobables. En nuestra propuesta didáctica no hemos diseñado tal problema, aunque sí dejamos como referencia el problema 3 de la secuencia propuesta en el libro [30], (ver figura 5.3).

Explicación del juego:

Se juega entre dos participantes. Cada jugador elige un caballo y marca con un color el número que lo identifica. No pueden tener el mismo caballo (si no se ponen de acuerdo, cada uno lanza un dado y elige primero quien obtenga el número mayor).



Se lanzan los dos dados y se suman los números que salen. El caballo cuyo número coincide con esa suma avanza un casillero. Gana la partida el jugador cuyo caballo llega primero a la meta. Cuando esto ocurre, la partida finaliza.

Figura 4.3: Enunciado de Carrera de Caballos

En la propuesta que realizan les autores se propone un trabajo exploratorio, que combina experimentación por parte de les estudiantes en torno a cuál es el caballo que más veces gana, y su relación con la probabilidad de ganar del mismo. En este problema, los caballos avanzan de un casillero, sobre la columna en la que se encuentran ubicados en el momento inicial. Consideramos que este problema puede ser una buena oportunidad para discutir con les estudiantes el cálculo de probabilidades en un espacio equiprobable, la puesta en relación de este cálculo con la probabilidad estimada frecuentísticamente.

Vemos una potencialidad, que no exploraremos en detalle en esta tesis, en torno a proponer para este juego una serie de items tendientes a la consideración de un torneo compuesto por una gran cantidad de partidos de este juego. Adecuadamente diseñados, estos items podrían alojar producciones matemáticas asociadas al cálculo de la probabilidad de ocurrencia de un gran número de eventos, de modo que la experiencia de trabajo del primer problema y de esto que esbozamos pueda funcionar como una apoyatura para la definición axiomática de probabilidad. Con estos dos problemas, creemos que le docente podrá proponer un trabajo en el aula específicamente dedicado al desarrollo de TP, donde podrá apoyarse en el trabajo sobre estos problemas para articular las diferentes definiciones y propiedades que puedan ser producidas a al finalizar los problemas.
Consideramos que la definición axiomática de espacio y medida de probabilidad puede ser puesta en juego, con la deducción, a partir de los axiomas, de las propiedades que se hayan formulado a partir de la definición frecuentista.

4.4.2. Problema 2

Para el abordaje de este problema que diseñamos, les estudiantes deberán haber construido técnicas y sentidos sólidos para el cálculo de probabilidades de eventos asociados a experimentos simples (como ser el arrojado de una moneda o de un dado), en los que la idea de probabilidad como cociente de casos favorables y casos posibles haya sido discutida y puesta en juego, y en los que se haya discutido el uso del número combinatorio para el cálculo de estas probabilidades.

Este problema está concebido con la intención global de trabajar sobre la propuesta de modelos probabilísticos en interrelación con una determinada experimentación, si bien ya el problema 1 avanza en esta dirección. En líneas anteriores, nos referimos a la noción de Modelo Estocástico como uno en el que, a propósito de la modelización matemática de fenómenos o problemas intra o extra-matemáticos, les estudiantes articulan un modelo probabilístico con una experimentación repetible para la producción de conocimiento sobre dicho problema. En el problema que proponemos, el modelo probabilístico está en función de describir el comportamiento de una familia de artefactos y la experimentación parte de la utilización (vía simulación) de algunos de ellos.
Más en particular, este problema busca abordar la noción de distribución binomial para el caso particular de parámetros $n, \frac{1}{2}$, no necesariamente apoyada en la noción abstracta de variable aleatoria, de modo que la posterior definición pueda apoyarse en las producciones anteriores.

Enunciado y Análisis del Problema 2

2) En este problema nos proponemos estudiar el comportamiento del tablero de Galton, mediante el simulador https://www.edumedia-sciences.com/es/media/905-maquina-de-galton

Antes de proponer el ítem a) y como parte de la devolución de este problema, le docente debería, en el gran grupo, explicitar algunos aspectos del tablero y del simulador, mostrando uno que tenga 4 filas, respectivamente con 1, 2, 3 y 4 lugares donde pueden rebotar las bolillas. Simular el tiro de varias bolillas con la pregunta de "¿qué ven que ocurre en el simulador?", puede ser un punto de partida para llegar a un consenso en torno a cómo funciona este artefacto. En particular, que en cada nivel la bolilla puede ir hacia la izquierda o hacia la derecha, que la cantidad de lugares donde puede caer es uno más que la cantidad de filas, y que el comportamiento es relativamente ideal (las bolillas siempre rebotan del mismo modo, si tiramos más de una no se tocan entre sí). También, las opciones que ofrece el simulador (variar la cantidad de filas y la cantidad de tiros). Consideramos que, si no surge explicitamente, le docente puede incorporar que la bolilla no tiene preferencia de ir hacia la derecha o a la izquierda. No obstante, es posible que les estudiantes asuman este supuesto de manera implícita, y en la puesta en común del ítem a) podría explicitarse. Consideramos necesario explicitar que esperamos que les estudiantes, eventualmente, numeren las columnas del 1 al 5. Según la conveniencia de este escrito, a veces nos referiremos a la numeración del 0 al 4 para mayor simpleza.

a) Supongamos que tenemos un tablero con 4 filas. Si tiran una bolilla, ¿cuál es la columna en la que es más probable que caiga? ¿Y la menos probable? Argumenten.

Una variable didáctica para considerar en este problema es si el lanzamiento del ítem a) es mediante la utilización o no del simulador por parte de les estudiantes. Si bien en los siguientes ítems se trabajará extensivamente con el simulador, en este es interesante que les estudiantes, a partir del proceso de devolución llevado adelante antes del lanzamiento del ítem a), puedan anticipar el comportamiento de la bolilla, para que luego la experiencia con la simulación permita contrastar con dichas anticipaciones. No obstante, es posible que para algunes estudiantes esta anticipación sin uso del simulador sea costosa, y que poder utilizar el simulador les permita construir representaciones personales sobre el funcionamiento del artefacto.
Según el dominio de técnicas combinatorias que se haya construido en el aula, este problema puede ser planteado originalmente con un tablero de 6 filas.
Este ítem, de resolución autónoma por parte de los pequeños grupos, la intención es que les estudiantes anticipen el comportamiento aleatorio de la bolilla en relación con los posibles recorridos que esta puede describir, y con los diferentes lugares en los que la bolilla puede caer. En el escenario equiprobable, que anticipamos que estudiará la mayoría -si no todos- de los grupos, todos los recorridos son igual de probables, aunque la cantidad de ellos que desemboca en cada lugar del final del tablero es diferente, por lo que la equiprobabilidad se pierde al estudiar el evento "cae en la columna i-ésima".
Fijada la cantidad de filas, digamos 4 en este caso, se puede modelizar el recorrido de la bolilla como una tira de 0's y 1's, en la que un 0 o un 1 en la posición j de la tira, representa que la bolilla, en la j-ésima fila, va hacia la izquierda o la derecha respectivamente. Así, si las columnas se numeran del 0 al 4, la cantidad de 1´s en la tira se corresponde con el número de columna en la que cae la bolilla (por ejemplo: una tira de 0's representa que la bolilla siempre ha ido a la izquierda, entonces caerá

en la columna 0). Así, la cantidad de caminos que llegan a la columna i-ésima es la cantidad de tiras que tienen i 1's en ella, y puede contarse mediante el número combinatorio $\binom{4}{i}$. Bajo la hipótesis de equiprobabilidad, todos los recorridos son igual de probables, y entonces la comparación numérica de los combinatorios permitirá la comparación probabilística de las columnas. No obstante, esta resolución matemáticamente correcta no es necesaria para el abordaje de este ítem en particular.
Anticipamos que la suposición de que las columnas de los extremos serán menos probables que las centrales estará disponible para les estudiantes, aunque posiblemente comparar cuáles de las centrales sean las más probables pueda resultar más complejo y ser fuente de dudas. En este sentido, los argumentos requerirán apelar a un modelo probabilístico, ya sea que este sea o no formulado explícitamente.
Por su parte, podría ser posible que algunes estudiantes consideren igual de probables todas las columnas, apelando a una idea de que "la bolilla no tiene preponderancia por ir hacia la izquierda o hacia la derecha". Consideramos que, si en la puesta en común esta idea no entra en conflicto, sí lo hará en el ítem siguiente, cuando se realice la simulación.
Que los lugares de los extremos sean poco probables puede ser sostenido con un argumento del tipo "alcanza con que una sola vez la bolilla cambie de dirección para que no caiga en ese lugar", mientras que para las centrales "hay muchos caminos que llevan a ellas". En una puesta en común, puede ser interesante preguntar cuántos son estos caminos para cada una de las cinco columnas, considerando que, por la simetría del problema, esto puede ser reducido al cálculo de tres. Posiblemente, algunos grupos se propongan calcularlas, lo cual para quienes aún no hayan construido cierto dominio combinatorio esto puede ser desafiante. Es necesario que le docente preste atención a las producciones a fin de detectar estas situaciones y poder abordarlas en una interacción directa con les estudiantes en cuestión.
La elección de una cantidad inicial de 4 filas permite que el cálculo combinatorio de los diferentes recorridos sea posible de ser hecho a mano, por ejemplo, con un diagrama de árbol o escribiendo la lista. Les estudiantes podrán proponer que el recorrido de la bolilla sea descripto (modelado) como una serie de cuatro resultados (izquierda o derecha), y que entonces cada recorrido puede ser escrito como una palabra que tenga cuatro letras, algunas de las cuales son letras D y otras son letras I. Con este modelo, la palabra IIII es la única que corresponde con caer en la columna de la izquierda, mientras que para cualquier otra (que no sea la columna de la derecha) hay más caminos. Para calcular cuántos, deberían comprender que la cantidad de letras I (o de letras D) determina cuál es la columna en la que el recorrido desemboca, y para hacerlo bien pueden recurrir a un argumento combinatorio como el que esbozamos más arriba, o a un diagrama de árbol.

Otra opción es que les estudiantes adopten una suerte de "modelo de cancelaciones", en el cual ir a la derecha suma una unidad, e ir a la izquierda resta una unidad. Este modelo no se encuentra plenamente adaptado. Si la numeración de las columnas es del 1 al 5, cosa que anticipamos, les estudiantes deberán reajustar el modelo, ya que, por ejemplo, el recorrido que desemboca en la columna 5 tiene asociado, en el esquema de cancelaciones, el número 4. Anticipamos que el estudio de la adaptación necesaria para este modelo requerirá de la intervención docente, y en este sentido lo consideramos una buena oportunidad para discutir en el aula.
Puede ser interesante que le docente proponga una instancia de puesta en común de las anticipaciones

de cada grupo, con el fin de que las diferentes respuestas y modelos propuestos puedan ser sometidas a un tratamiento colectivo. En este tratamiento, será particularmente interesante discutir en el aula la pertinencia de maneras de responder que apelen a diferentes maneras de describir la aleatoriedad del fenómeno.

Un objetivo de esta puesta en común es arribar al consenso de que los lugares más probables son los centrales, y que los más improbables son los de los extremos. Nos parece importante que en esta puesta en común se puedan abordar las diferentes respuestas de les estudiantes, incluyendo en ellas los discursos que las sostienen. Vemos un potencial en que le docente pueda fomentar que todes les estudiantes tengan la posibilidad de acotar, preguntar y objetar a las resoluciones de sus compañeres aquellas cuestiones que no entiendan. Otra componente importante de esta puesta en común será la articulación de las soluciones, por ejemplo, en el caso de que algunes estudiantes que hayan apelado a un diagrama de árbol y otres a utilizar un argumento combinatorio. Es posible que les estudiantes que realicen el diagrama de árbol no logren concluir la relación entre la cantidad de veces que va a la izquierda (o derecha) y la columna en la que desemboca. En cambio, quienes apelen al uso del número combinatorio posiblemente hayan arribado a una versión más completa de esta idea. Le docente, en su labor de gestor y regulador de estas interacciones, podrá apoyarse en las ideas de unes y otres para perseguir el objetivo de que estas vayan evolucionando en el desarrollo del problema.

Para el abordaje del siguiente ítem, será necesario que le docente discuta antes con les estudiantes la misma pregunta con un tablero de 5 y de 6 filas. En este caso, esperamos que les estudiantes conjeturen que, nuevamente, las columnas de los extremos serán menos probables, y que la más probable será la del medio (anticipamos a su vez que será necesario discutir que, en el caso de 5 filas, hay dos columnas más probables, cosa que podría pasar inadvertida).
En esta interacción será necesario realizar el cálculo, entre todes, de la probabilidad de caer en los extremos, y la probabilidad de caer en el lugar central, de modo que funcionen como un soporte para esas afirmaciones. La realización de este cálculo para el caso de 5 filas podrá todavía apoyarse en un diagrama de árbol, aunque ya para el caso de 6 le docente deberá priorizar escrituras y razonamientos que involucren el uso del número combinatorio.
Si bien consideramos deseable que calculen, para el caso de 6 filas, la probabilidad de caer en cada una de las columnas -y no solo de las más y menos probables-, optamos por proponerlo en el ítem c). Nos imaginamos que, en el pizarrón, podría quedar desplegada una figura como la siguiente, al lado de otra análoga para el caso de 5 filas.

b) Si tiran 1000 bolillas en un tablero de 6 filas, ¿qué esperan que pase? Realicen el experimento y comparen con sus anticipaciones.

La pregunta por 1000 bolillas busca ponerse en diálogo con lo trabajado en el problema 1, en torno

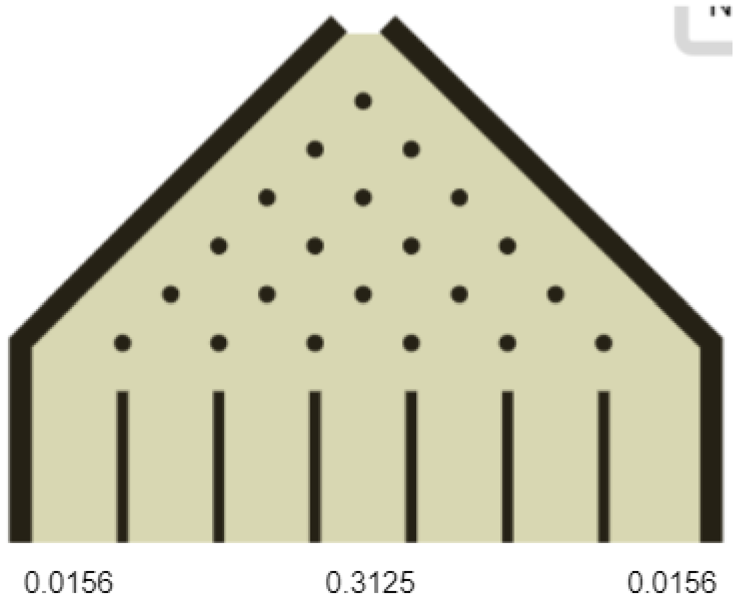


Figura 4.4: Diagrama de un tablero de 6 filas, que podría quedar desplegado en el pizarrón antes del ítem b)

a que la cantidad esperada de bolillas en una determinada columna es un número cercano a la probabilidad calculada en la puesta en común anterior, multiplicado por 1000.

Las respuestas de les estudiantes a la pregunta de "qué esperan que pase" pueden ser diversas, y nos interesa que puedan apoyarse en las relaciones provisorias a las que llegaron en el ítem anterior, incluyendo la puesta en común. Si bien consideramos que la escritura numérica en el pizarrón puede fomentar que algunes estudiantes se apoyen en ella para anticipar qué pasará con la realización 1000 veces del experimento, podría ser que otres se apoyen exclusivamente en las relaciones cualitativas a las que llegaron en el ítem anterior, para afirmar que en la columna central caerán más bolillas que en el resto, y que en las columnas de los extremos habrá menos bolillas que en las demás. Nos parece importante que la puesta en común recupere la pertinencia de ambas propuestas. Asimismo, es posible que haya estudiantes que puedan ser exhaustives en su respuesta, proponiéndose anticipar qué pasará con todas las columnas. En este caso, deberán calcular las probabilidades y compararlas entre sí, a fin de poder anticipar experimentalmente, ya sea cualitativamente (cuáles columnas tendrán más bolillas que otras) o cuantitativamente (una estimación de cuántas bolillas habrá en cada columna). Anticipamos que las respuestas cuantitativas serán diferentes (a causa de que cada grupo tendrá diferentes resultados experimentales), y que se apoyarán en la similitud, para valores grandes, entre la proporción de bolillas que cayeron en la columna correspondiente, y la probabilidad de que caigan en ella. Sería interesante que le docente deje escrito en el pizarrón, al retomar estas producciones, que

$$\frac{\text{Cantidad de bolillas que cayeron en la columna i-ésima}}{\text{Cantidad de bolillas tiradas}} \sim \mathbb{P}(\text{Caer en el la columna i-ésima}) \quad (4.1)$$

Explicitando que el parecido entre uno y otro número responde al mismo fenómeno que se trabajó en el problema 1, y diciendo que el valor de la izquierda es una estimación experimental del valor de la derecha. Asimismo, y apoyándose en esta interacción colectiva, al no haber discutido en el ítem anterior la probabilidad de caer en cada una de las otras columnas, les estudiantes pueden abordar autónomamente el siguiente ítem, que a su vez permite que les estudiantes que no hayan realizado una mirada cuantitativa, tengan oportunidad de hacerlo. Como dijimos, es posible que algunes estudiantes ya hayan realizado una caracterización del fenómeno aleatorio en cuestión, por lo cual le docente podría optar por, a partir de sus producciones -y de su necesaria puesta en diálogo con las otras producciones-, abordar directamente el ítem c) colectivamente.

c) Estimar y calcular, para un tablero de 6 filas, la probabilidad de caer cada una de las columnas.

Consideramos que la realización de este ítem requerirá por parte de les estudiantes, en primera instancia, de la distinción entre estimar y calcular. Le docente podrá aclarar oralmente, si no hubiera sido construida esta distinción en las experiencias de trabajo anteriores, que el término “estimar” remite a realizar una aproximación mediante datos -y que en el caso particular de estimar probabilidades tiene como herramienta privilegiada la utilización de las frecuencias relativas-, mientras que “calcular” requiere de la propuesta de un valor exacto basado en un modelo matemático establecido mediante un conjunto de supuestos, apoyándose en la puesta en común del ítem anterior, para lo cual haber escrito la ecuación (4.1) será conveniente.
Para la estimación, esperamos que les estudiantes dividan por 1000 la cantidad de bolillas de cada una de las columnas (empezando por alguna de ellas). La elección del número 1000 apunta a posibilitar un tratamiento numérico de las cantidades experimentales sin necesidad de hacer cada una de las cuentas en una calculadora.
Para el cálculo, será necesario que le docente preste atención a les estudiantes que encuentren dificultades para realizar el conteo, ya sea a partir de realizar un diagrama de árbol, o de manipular números combinatorios. No obstante, anticipamos que algunas de estas dificultades habrán tenido que surgir durante la puesta en común anterior, ya que el cálculo de la probabilidad de caer en la columna central ya apela a estos conocimientos.

En la puesta en común pretendemos que le docente pueda recopilar las diferentes soluciones. Si hubiera grupos que llegaron a resultados erróneos, sería interesante tomarlos en la puesta en común para discutir en torno a ellos. Los errores que anticipamos estarán relacionados con un conteo incorrecto de los recorridos que llevan a alguna de las columnas, y vemos fértil que la discusión colectiva de estos cálculos pueda apoyarse, ya sea en argumentos matemáticos en torno al conteo, como a la relación entre el valor experimental obtenido y el valor teórico propuesto por elles (podría ser necesario realizar la experimentación con 10000 bolillas para que este argumento sea más eficaz). Consideramos que en el pizarrón podría completarse el esquema anterior, incorporando los datos de la experimentación de 1000 tiradas realizada por le docente, quedando posiblemente algo como lo siguiente, siendo necesario incluir los cálculos realizados y las comparación con la estimación correspondiente:

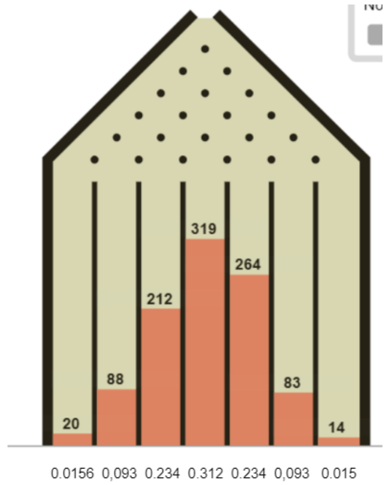


Figura 4.5: Diagrama que podría quedar desplegado en el pizarrón antes del ítem c)

En la instancia de institucionalización de este problema, le docente podrá dar una definición provisoria de variable aleatoria, entendida como valor númerico asociado al resultado de un experimento cuyo resultado es impredecible, dejando para el final del problema 3 la definición abstracta como función medible del espacio muestral a $\mathbb{R}$.

▲

Este mismo problema puede ser retomado en el estudio del Teorema Central del Límite, siendo la maquina de Galton uno de los instrumentos más populares empeados para la presentación de la *campana de Gauss* para modelar suma de variables aleatorias simétricas.

4.4.3. Problema 3

Para el abordaje del siguiente problema, que nos parece importante que se realice temporalmente cerca del anterior, se requiere del conocimiento de la noción de independencia. En el capítulo anterior esbozamos unas discusiones y consideraciones que consideramos fértiles de tener en cuenta para el diseño de uno o más problemas que avancen en esa dirección. A estos fines, podría diseñarse, por ejemplo, un problema que retome un torneo de carreras de caballos como el mencionado en líneas anteriores, en el que se ponga en relación la técnica de cálculo de la probabilidad de ganar K veces seguidas una carrera con el caballo 7, y la probabilidad de ganarla individualmente, para conjeturar la regla de multiplicatividad en este escenario. Esto podría ser una apoyatura para proponer experiencias de trabajo específicas para el abordaje de la noción abstracta de independencia y su relación con la probabilidad condicional.

El objetivo de este problema es construir sentidos en torno a una distribución Binomial de parametro p usando nuevamente un tablero de Galton donde ahora los rebotes no ocurren de manera simétrica. Se espera que puedan, colectivamente, arribar a la fórmula de la probabilidad puntual y estudiar de que manera esta varía a lo largo de su rango, identificando relaciones de crecimiento y decrecimiento. También, dar los primeros pasos en torno a la comparación de estimadores. También, la posibilidad

de experimentar con el tablero permite una vez mas calcular frecuencias relativas y compararlas con las probabilidades calculadas bajo el modelo, evidenciando nuevamente la relación empírica existente entre la frecuencia relativa y al probabilidad. En este problema, les estudiantes modelarán un tablero de Galton con p, la probabilidad de ir hacia la derecha, es $3/4$ y con 6 filas. Vemos importante que el tablero se encuentre proyectado en el pizarrón, de modo que mientras avanza el proceso de devolución se vean bolillas caer, y que vayan haciéndolo de a una, de modo que progresivamente se vaya mostrando en el simulador la cantidad que ha caído en cada columna.
Sería interesante que le docente sostenga algunos intercambios con les estudiantes a fin de llegar a un consenso respecto del comportamiento del tablero nuevo y su diferencia con el del problema 2. Para el abordaje de este problema, les estudiantes tendrán elementos de su trabajo anterior en los que podrán apoyarse (la manera de modelizar los recorridos). En este problema, salvo cuando $p = 1/2$, estamos en un modelo que deja de ser equiprobable y, por lo tanto, la división de casos favorables por casos posibles para el cálculo de probabilidad es incorrecta.

Enunciado y Análisis del Problema 3

a) Consideren ahora el Tablero de Galton -utilizando el simulador `www.mathsisfun.com/data/quincunx.html` **en el que, debido a una causa desconocida, las bolillas, al rebotar, tienen una determinada preferencia a irse hacia la derecha en vez de hacia la izquierda, independientemente de en cuál nodo reboten. Ajusten, en el simulador, esta preferencia, ubicándolo en el valor $\frac{3}{4}$. En este nuevo escenario, ¿cuál es la columna más probable? ¿Y la menos probable? Para realizar esta tarea, numeren las columnas del 0 al 6**

En lo que sigue, analizamos posibles resoluciones de les estudiantes. Para el cálculo de la cantidad de caminos que llevan a una columna determinada, les estudiantes podrían apoyarse parcialmente en los razonamientos combinatorios desplegados en el problema anterior. La novedad que se juega en este problema es la de la preponderancia, que les estudiantes tomarán o no en cuenta en sus producciones de manera diversa. Avanzamos en esta dirección:
Algunes de estes estudiantes podrían, erróneamente, concluir que la columna más probable es la del medio, omitiendo la nueva información sobre el comportamiento del tablero. Otres, poniendo el foco sólo en la preponderancia de ir hacia la derecha, podrían concluir que la columna más probable es la de la derecha puesto que el recorrido que va hacia dicha columna es el que más en cuenta toma el comportamiento de las bolillas.
Otra resolución posible es que les estudiantes planteen que, del total de veces que la bolilla rebota, $\frac{3}{4}$ veces irá a la derecha y $\frac{1}{4}$ parte de las veces irá a la izquierda. Como hay 6 ocasiones en los que la bolilla rebota, podrían proponerse calcular el resultado de la multiplicación $6.\frac{3}{4}$. El resultado de esta cuenta es $4,5$, por lo que algunes estudiantes podrían responder que la columna más probable es la 4 o la 5, o que este razonamiento no les permite inclinarse por ninguna de las dos, y que entonces ambas son igual de probables.
Todes les estudiantes podrán apelar al uso del simulador para, ya sea responder apoyades en sus resultados, o bien contrastar sus respuestas con los resultados de la simulación, porque el valor de p del

simulador estará determinado en $p = \frac{3}{4}$.
Una respuesta correcta para la columna menos probable podría apelar a combinar la preponderancia y la cantidad de caminos, observando que hay un solo camino que desemboca en la columna que está más a la izquierda, y este camino debe ser el más improbable porque la bolilla preponderantemente va hacia la derecha.
Finalmente, algunes estudiantes podrían apelar al uso explícito del valor $\frac{3}{4}$ en el cálculo de la probabilidad de caer en cada columna. Para ello, es preciso sumar la probabilidad de cada uno de los caminos que lleva a dicha columna. Estos caminos pueden ser identificados con una palabra de seis letras, D e I, con i letras D, y $6-i$ letras I. Por lo cual, dado un camino que lleva a la columna i, y gracias a la independencia de los rebotes, resulta que su probabilidad es $\left(\frac{3}{4}\right)^i . \left(\frac{1}{4}\right)^{6-i}$. Esto muestra, además, que todos los caminos que llevan a dicha columna tienen igual probabilidad, por lo que para calcular la probabilidad de caer en la columna i, basta ahora con contar -del mismo modo que en problema anterior- cuántos caminos conducen a dicha columna- Esa cantidad es $\binom{6}{i}$. La comparación numérica de la probabilidad de caer en cada columna permitirá que les estudiantes respondan correctamente. Algunas resoluciones que apelen a estas ideas pueden presentarse incompletas (por ejemplo, algunes estudiantes podrían lograr calcular la probabilidad de ir a la columna de la derecha o de la izquierda, pero podrían encontrar alguna complicación a la hora de contemplar caminos que incluyan rebotes tanto hacia la izquierda como hacia la derecha).

Consideramos pertinente la propuesta de una puesta en común para abordar las producciones de cada grupo. En este caso, anticipamos que habrá grupos que quizás no lograron producir una respuesta completa, y en este sentido será interesante someterlo a un tratamiento colectivo.
Le docente, en la puesta en común, tendrá potencialmente cinco respuestas diferentes por parte de les estudiantes. En algunas, les estudiantes podrían:

- Apelar a la noción de independencia y a la descripción de los recorridos como palabras con letras I y D, para calcular la probabilidad de caer en una o más columnas. Según el grado de exhaustividad de estas resoluciones, podrán responder o no cuál es la columna más y menos probable.
- Concluir que la más probable es la de la columna 6, solo tomando en cuenta la preponderancia de ir hacia la derecha, y apoyarse en esta idea para afirmar que la menos probable es la columna 0.
- Privilegiar la cantidad de caminos pero no la preponderancia, concluyendo que la más probable es la columna 3, y que las menos probables son la 0 y la 6.
- Concluir que la columna 0 es la menos probable, tomando en cuenta la preponderancia y la cantidad de caminos, pero sin responder cuál es la columna más probable.
- Responder correctamente apoyades en la simulación.

Nos interesa observar que, con excepción de la primera y cuarta solución, las otras resoluciones pueden no apelar en ningún momento a una idea de independencia en el sentido probabilístico.

Aunque sí existe una potencialidad que le docente puede aprovechar, que es la que da la comparación entre las respuestas. Tanto la segunda como la tercera son erróneas, lo cual la quinta logrará mostrar. Entonces la pregunta por la razón por la que las respuestas son erróneas, aunque puedan resultar convincentes, puede permitir a le docente un intercambio de ideas con les estudiantes a partir de lograr un posicionamiento reflexivo respecto de sus producciones. En este escenario, consideramos que habrá buenas condiciones para que le docente proponga relacionar la preponderancia por ir hacia la derecha, presente en la segunda resolución, con la cantidad de caminos, presente en la tercera. Si en el aula hubo estudiantes que realizaron la primera resolución, entonces le docente podrá apelar al análisis de esta para discutir esa relación.

Ahora bien: podría pasar que ningún grupo de estudiantes hayan intentado calcular la probabilidad de caer en cada columna. Esto nos abre dos escenarios, en función de los cuales proponemos diferentes acciones docentes:

Escenario 1 Si no hubo producciones que den una manera de calcular la probabilidad de caer en cada columna, sería interesante que le docente deje explícito en el pizarrón las conjeturas y cuestiones que surgieron, especialmente las siguientes:

(I) La columna más probable es la quinta, y la más improbable es la primera.

(II) La bolilla tiene mayor preponderancia a ir hacia la derecha, pero la séptima columna, correspondiente a $i = 6$, no es la más probable.

(III) La cantidad de recorridos que desembocan en cada columna es la misma que en el problema 2.

(IV) La probabilidad de ir hacia la derecha es la misma para todos los nodos.

Con esto en mente, podría proponer la siguiente pregunta **para responder entre todes:**
¿Cuál es exactamente la probabilidad de caer en la columna 6? ¿Y en la 4?
Esta pregunta vuelve necesario distinguir que, si bien dónde rebota la bolilla depende fuertemente de dónde rebotó antes, *la dirección en la que rebota en uno sí es independiente de la dirección en la que rebota en el siguiente*. Es decir, las direcciones sorteadas en los diferentes nodos son independientes entre si. Esto puede permitir que, si les estudiantes logran caracterizar al recorrido en términos de eventos independientes (donde el evento independiente en cuestión es “ir hacia la izquierda o hacia la derecha”), entonces podrán dar explícitamente la probabilidad, anticipamos, sin mayores complicaciones. Pero depende de que les estudiantes modelicen la situación mediante la idea de independencia. Si esto no ocurre, será necesario que le docente lo explicite y realicen los cálculos correspondientes. En este caso, proponemos que le docente gestione la discusión en torno al cálculo de la probabilidad de caer en la columna 6, dejando en el pizarrón una escritura como la siguiente:

$$\text{si } p = \frac{3}{4}, \text{ entonces } \mathbb{P}(\text{Caer en la columna 6}) = \left(\frac{3}{4}\right)^6$$

y deje un espacio de trabajo autónomo para el cálculo de la probabilidad de caer en la columna 4.

Será interesante que en la puesta en común posterior a esta instancia, le docente pueda incorporar una escritura como la siguiente

$$\mathbb{P}(\text{Caer en la columna 4}) = \binom{6}{4}\left(\frac{3}{4}\right)^4\left(\frac{1}{4}\right)^2$$

Luego de ello, quedaría para abordar colectivamente la respuesta de cuál es la columna más probable, computando la probabilidad de caer en las columnas restantes y comparando los valores numéricos obtenidos.

Escenario 2 Si hubo producciones que empleen explícitamente el "p" para el cálculo de probabilidades, le docente puede discutir los supuestos que subyacen a ese cálculo (la independencia de la dirección de los rebotes, y que todos los caminos que que conducen a una misma columna tienen igual probabilidad), con especial atención a aquelles estudiantes que no hayan apelado a este tipo de razonamientos, con la intención de poner en diálogo sus ideas.

En función de lo trabajado en ambos escenarios, se habrá arribado a la probabilidad de caer en cada columna, y concluido cuál de las columnas es la más y menos probable. Luego, proponemos que le docente lance el item b) de resolución colectiva.

b) Dado un tablero de 6 filas en el que las bolillas tienen probabilidad $p = \frac{3}{4}$ de ir hacia la derecha en cada uno de los nodos, realizar un gráfico que relacione el número de columna con la probabilidad de caer en dicha columna.

Al recuperar el valor $\frac{3}{4}$, graficarlo en este caso podría ser accesible ya que se habrá calculado la probabilidad de caer en cada una de las columnas. Con este trabajo realizado, le docente puede instalar la pregunta por cómo será el gráfico si el valor de p es diferente. Proponemos que le docente aborde esta tarea de para $p = \frac{k}{7}$, donde k es un entero entre 0 y 6. En cada gráfíco, la columna más probable será la k—ésima. Anticipamos que con lo discutido hasta el momento, le docente podría apoyarse para escribir, en función de un p cualquiera

$$\mathbb{P}(\text{caer en la columna i}) = \binom{6}{i}.p^i.(1-p)^{6-i} \quad \text{para} \quad 0 \leq i \leq 6$$

c) Supongamos que tenemos un nuevo tablero en el que no se conoce la probabilidad de que la bolilla rebote hacia la derecha. Al tirar una sola bolilla esta cae en la columna 4. ¿Cuál estimarían que es el valor de $p = \mathbb{P}$(probabilidad de que la bolilla rebote hacia la derecha)? ¿Y cómo sería si cayera en la columna 3?

Para responder, les estudiantes deberán movilizar la idea de que en este modelo, caer en la columna 4 equivale a que, cuatro de seis veces, la bolilla haya ido a la derecha. Retomando la relación entre frecuencia relativa y probabilidad que fue explorada en los ejercicios anteriores, es posible que les estudiantes estimen entonces la probabilidad p con el valor $\frac{4}{6}$, para la columna 4, y con $\frac{3}{6}$ para la columna 3. Más generalmente, si X denota la columna donde cae una bolilla, podemos estimar p con $X/6$.

Este item busca instalar en el aula una pregunta que, además de ponerse en diálogo con el item d), busca discutir una estrategia que puede seguir siendo discutida en problemas posteriores: la de, antes de proponer un estimador basado en grandes números, primero pensar en un estimador basado en una única realización del experimento.
Anticipamos que algunes estudiantes querrán apelar al uso del simulador para grandes números, ya que como estrategia esta ya estará disponible. En este sentido, le docente deberá reponer que eso será objeto del siguiente item, en el que se les da a les estudiantes un simulador de un tablero con 6 filas donde ahora no conocen el valor de p. En lo que sigue, utilizamos $p = \frac{2}{3}$.

d) Estimar el valor de

$$p = \mathbb{P}(\textbf{probabilidad de que la bolilla rebote hacia la derecha})$$

a partir de los datos experimentales obtenidos en el simulador.

Para realizar este item, les estudiantes pueden apelar a, al menos, dos estimaciones diferentes:

- A partir de lo realizado en el ítem anterior, promediar las estimaciones obtenidas utilizando la columna correspondiente a cada bolilla. Matemáticamente hablando, si X_i denota la columna donde cae la i-ésima bolilla, tenemos que promediar los valores $X_i/6$. Luego, qla estimación para p está dada por $n^{-1}\sum_{i=1}^{n} X_i/6$, donde n denota la cantidad de bolillas utilizadas en el experimento.
- A partir de lo realizado en el item b), considerar, para una columna dada, que la proporción de bolillas que cayó en ella es una aproximació de la probabilidad de caer en ella, y luego intentar despejar el valor de p a partir de conocer la probabilidad de caer en dicha columna.

Observemos que, con relación a este último estimador, despejar el valor de p a partir de la columna 4 no es posible sin ayuda de un software que aproxime la solución, ya que implica resolver la ecuación $q = 15p^4(1-p)^2$, donde q es la frecuencia relativa asociada a la columna $i = 4$ calculada con los valores que el simulador haya arrojado. Esta dificultad, esperamos, podrá funcionar para que algunes estudiantes se propongan generalizar el procedimiento propuesto para el item b), o para preguntarse si el despeje podría ser posible a partir de considerar otras columnas. Esto último podría llevar a la consideración de la columna 6 o la 0, para la cual el despeje sí es posible. Anticipamos que, como experimentalmente habrá más bolillas en la 6 que en la 0, la mayoría de les estudiantes apelarán a la 6.
Nos parece importante proponer una puesta en común donde le docente pueda recuperar las diferentes producciones de les estudiantes, poniendo en discusión cuáles estimadores fueron utilizados, y cuáles las estimaciones obtenidas. Esta puesta en común puede ser una oportunidad para dar una primera discusión en torno a la diversidad de estimadores y la posibilidad de discutir criterios para decidirse por unos u otros. Sostener estas discusiones no es posible a partir sólo de datos particulares, sino que

será interesante discutir qué resultados arroja cada uno para una cantidad diversa de repeticiones. Nos parece fértil apelar a la una simulación de los mismos, y graficar histogramas correspondientes a diferentes métodos de estimación, para cantidades diversas de repeticiones. Al hacerlo, se podrá atrapar la diferencia en la dispersión de las estimaciones con uno u otro procedimiento.

Con esto en mente, y en perspectiva de comenzar a articular con el problema 2, nos parece interesante que le docente indague en cuáles variables aleatorias (con el sentido provisorio explicitado en el problema 2) fueron estudiadas en el problema 3, y en el problema 2, apuntando a considerar en un primer momento, a la variable $Y =$ Número de columna en el que cae la bolilla en un tablero de N filas. Creemos que le docente puede incorporar esa escritura al pizarrón, y preguntar qué características de esa variable fueron estudiadas, con la intención de relacionarlo con la noción, que será definida in situ, de rango y función de probabilidad puntual. Los elementos de la puesta en común que esbozamos en líneas anteriores serán puntos de apoyo para esta discusión. Observamos que será necesario, para referir al trabajo realizado con el tablero de Galton, discutir en el aula que la posición desde la que cae la bolilla es:

- Alineada con la columna del medio, cuando N es par.
- Alineada con la barra que separa a las dos columnas del medio, cuando N es impar.

Con esto en mente, puede proponerse cálcular el rango de Y y de su función de probabilidad puntual. Esta puede ser una buena oportunidad proponer una primera definición de variable aleatoria binomial, a partir de la idea provisoria de variable aleatoria, y caracterizando a la binomial por su rango y su función de probabilidad puntual.
También, le docente podrá discutir qué es un histograma, a fin de tenerlo disponible para futuros problemas.

▲

A partir de aquí, le docente podrá proponer -partiendo de esta puesta en relación de los problemas 2 y 3- una instancia de institucionalización en la que ciertos objetos o relaciones cobren un estatus dentro de la construcción de TP. En particular, consideramos que es una buena oportunidad para que le docente formule en general la noción de rango y de función de probabilidad puntual, y aborde la noción formal de variable aleatoria como función del espacio muestral en $\mathbb{R}$, y defina mediante estos elementos la variable aleatoria binomial.

En perspectiva de realizar una revisión global de los elementos tecnológico-teóricos que se abordaron desde el primer problema y hasta el tercero, así como de las técnicas-tareas que se desarrollaron, consideramos pertinente dar un espacio de reflexión en torno al rol de la experimentación en los problemas. Esto permitirá reflexionar sobre lo abordado en el problema 1, en torno a la definición frecuentista de la probabilidad, y a la existencia de espacios muestrales no equiprobables, así como descontextualizar las

técnicas empleadas para el cálculo de probabilidades de eventos, ya sea en un espacio muestral equiprobable (casos favorables/casos posibles) y no equiprobable (descomposición de eventos en unión de eventos elementales). En esta misma dirección, se puede retomar la idea de independencia presente en el problema 3 y poner en común también lo construido a ese respecto. El lenguaje de la TP podrá utilizarse plenamente en este contexto para la expresión de las relaciones y propiedades correspondientes.
Luego de estos tres problemas, y las discusiones posteriores, consideramos que en el desarrollo de la materia será necesario realizar un abordaje de diferentes asuntos, que esbozaremos a continuación, para los cuales es necesario diseñar problemas que permitan estudiarlos, así como articular su puesta en aula con sucesivas instancias de institucionalización, como hemos esbozado hasta este momento. Concretamente, será necesario que les estudiantes hayan construido, a partir de experiencias de Modelización Estocástica, conocimientos en torno a:

- Diferentes variables discretas, ya sea con distribuciones conocidas (geométrica, hipergeométrica, binomial negativa, Poisson) como con distribuciones desconocidas pero calculables, de rango finito y numerable.

- Operaciones entre variables aleatorias discretas, en particular en torno a la suma, cálculo de las funciones de probabilidad puntual de las variables aleatorias definidas por dichas operaciones, modelización de problemas cuyos modelos involucren operaciones entre variables aleatorias. Además, la noción de independencia de variables aleatorias -será necesario para este abordaje una instancia de trabajo con vectores aleatorios bidimensionales-.

- Esperanza y varianza. Sentidos matemáticos y propiedades de la esperanza y la varianza del promedio de variables aleatorias, en particular el caso iid. Desigualdad de Markov y Tchebychev.

Por cuestiones de extensión de este trabajo, omitimos diseñar y analizar problemas que aborden estos asuntos. Con este mismo motivo, para los próximos dos problemas que siguen (y que completan la propuesta didáctica) delinearemos las cuestiones generales de su análisis didáctico-matemático.

4.4.4. Problema 4

Este problema tiene la intención de, a partir de las herramientas de Modelización Estocástica construidas hasta el momento, abordar una instancia de trabajo que permita, a posteriori del trabajo con este problema, formular y demostrar la Ley Débil de los Grandes Números para variables aleatorias con segundo momento finito.
Este problema es una adaptación del problema propuesto por Guy Brousseau (2002, [15]), en el que el objetivo es que les estudiantes estimen la composición de una urna que tiene una cantidad conocida de bolillas, algunas blancas y otras negras, pero no se conoce cuántas. Concebimos que le docente puede proponer este problema, ya sea dando a les estudiantes acceso a un simulador que permita repetir el experimento "sacar una bolilla, anotar si es blanca o negra, y devolverla" tantas veces como quieran, o

construir materialmente[7] una urna. En nuestro problema, la composición de la urna es 35 blancas y 65 negras.

Enunciado y Análisis del Problema 4

Se tiene una urna con 100 bolillas, algunas blancas y otras negras. Se quiere estimar la composición de la urna, que es desconocida. Para ello, pueden, de a una vez, sacar una bolilla, anotar el color, y devolverla, tantas veces como quieran.
a) Estimen la proporción de bolillas blancas, explicando el procedimiento que utilizaron.

En este ítem, bien abierto, anticipamos que les estudiantes propondrán la realización del experimento "sacar una bolilla" (o su análogo, "ver una bolilla", en el caso de que tengan la botella en físico) una determinada cantidad de veces, y contar cuántas de esas veces la bolilla salió blanca. Consideramos que en cada grupo podrá haber una variedad de experimentos propuestos. En lo que sigue, nos restringimos a los siguientes posibles experimentos:

- (de tipo 1) Repetir N veces el experimento "sacar una bolilla", y promediar.
- (de tipo 2) Repetir N veces el experimento "sacar K veces una bolilla, y promediar", y luego tomar promedio.

Consideramos necesario que haya un tiempo de exploración de las urnas/simuladores por parte de cada grupo, que utilizará para estimar uno de estos tipos de experimento, incluyendo en esta elección un N particular de ese grupo. Anticipamos que la gran cantidad de negras hará que, ante la repetición del experimento, les estudiantes conjeturen que habrá más negras que blancas. Si fuera el caso que algún grupo realiza varias repeticiones (por ejemplo, más de 10) y no salió ninguna vez una blanca, le docente podrá aclarar que en la urna hay blancas y negras.
Cada experimento propuesto dará diferentes resultados, y consideramos fértil que le docente comience la puesta en común preguntando primero, grupo por grupo, cuál es el valor que proponen para la proporción de blancas. Para realizar la estimación, anticipamos que algunes estudiantes dirán el promedio experimental obtenido, y otres el mismo redondeado.
Ante la gran variedad de respuestas, será interesante que le docente indague en cómo cada grupo realizó la estimación. Algunes estudiantes podrían adjudicar la multiplicidad de respuestas a la diferencia en la metodología empleada. A partir de esto, será necesario que le docente discuta con todes les estudiantes algunos asuntos:

1. Cómo los experimentos de tipo 2 pueden verse como experimentos de tipo 1.
2. Cómo expresar, en términos de variables aleatorias, el método de estimación propuesto (por ejemplo, un experimento de tipo 1 puede plantearse como un promedio de N variables Bernoulli independientes cuyo parámetro es el valor a estimar). Será necesario, al respecto de esto, resaltar

[7] Una manera de hacer esto es utilizando botellas de plástico de tapa transparente, cuyo cuerpo ha sido pintado con alguna pintura opaca. Una vez puestas adentro las bolillas y con la tapa cerrada, les estudiantes solo podrán ver bolillas al poner boca abajo la botella, y ver la bolilla a través de la tapa.

que esta expresión tomará en cuenta el N del grupo, y observar que algunes estudiantes -y otres no- apelarán al redondeo del valor obtenido en la experimentación. Sugerimos que esto último sea expresado coloquialmente.

3. Cuál de todas las estimaciones elegirían. Trás haber discutido que ambas propuestas son equivalentes, lo único que puede variar es la cantidad total de repeticiones realizadas para el calculo de la estimación. En lo que sigue, entra en juego la importancia de esta cantidad.

Esta puesta en común puede ser un buen insumo que les estudiantes tengan en cuenta para abordar el siguiente ítem:

b) Acotar inferiormente (de manera no trivial) la probabilidad de que, realizando su procedimiento, la diferencia entre el valor estimado y el valor de p no exceda las 0,005 unidades.

Al lanzar este ítem, le docente podría apuntar a consensuar con les estudiantes que el evento en cuestión puede expresarse en términos de $|\frac{\sum_{i=1}^{N} X_i}{N} - p| \leq 0,005$, donde X_i vale 1 si la i-ésima extracción es blanca y cero caso contrario. Recordemos que el valor de N dependerá de cada grupo. Observemos que, lo que se pide calcular, puede ser expresado también como

$$\mathbb{P}\left(Np - 0,005N \leq \sum_{i=1}^{N} X_i \leq 0,005N + Np\right).$$

La probabilidad depende de N y de p. Si un grupo eligió $N \leq 100$, entonces $0,005N < 0,5$ y luego la probabilidad pedida coincide con $\mathbb{P}(\sum_{i=1}^{N} X_i = k(N,p))$ donde $k(N,p) = [Np]$ si $Np - [Np] \leq 0{,}5$, y $k(N,p) = [Np] + 1$ en caso contrario. Les estudiantes, por lo trabajado en los problemas 2 y 3, podrán concluir que esta probabilidad será baja. Algunes estudiantes podrían querer modificar su N para mejorar la estimación, y este movimiento podría posibilitar que postulen que, tomando N cada vez más grande, entonces $\mathbb{P}\left(|\frac{\sum_{i=1}^{N} X_i}{N} - p| \leq 0,005\right)$ aumenta, e incluso algunes estudiantes podrían conjeturar que tiende a 1.

Otres estudiantes pueden apelar al uso de la desigualdad de Bienaymé-Tchebychev, para la cual ocurre un fenómeno similar en términos de que, para N menor a 1000, la cota no aporta valores significativos, pero sí para N mayor.

Otres estudiantes, por su parte, podrían encontrar complejo abordar cualquiera de estos dos asuntos. Nos parece interesante que le docente pueda estar pendiente durante el tiempo de trabajo autónomo de les estudiantes, para evaluar si puede realizar intervenciones que permitan que les estudiantes avancen en alguna de estas direcciones.

La puesta en común nos parece central para que, en el pizarrón, puedan desplegarse los argumentos utilizados, de los cuales algunos estarán incompletos o tendrán errores, y en este sentido, como siempre, reflexionar colectivamente sobre estos asuntos podrá ser un punto de partida para movilizar ideas matemáticas interesantes, tanto para les estudiantes productores como para les estudiantes que

discutan, objeten o realicen preguntas en torno a las producciones de sus compañeres.
Para quienes haya tomado un $N < 100$ y hayan intentado calcular la probabilidad de manera directa, le docente puede proponer una cota partir de graficar en un software la función

$$f(p) = \mathbb{P}\left(\sum_{i=1}^{N} X_i = k(N,p)\right) = \frac{N!}{[Np]!(N-[Np])!} p^{[Np]}(1-p)^{(N-[Np])}$$

Incorporamos un gráfico aproximado de la función, cuyo estudio podrá ser un punto de apoyo para, en el contexto del trabajo colectivo, que les estudiantes estimen que la probabilidad pedida con $N = 100$ será mayor a 0.09, independientemente de cual es el valor de p:

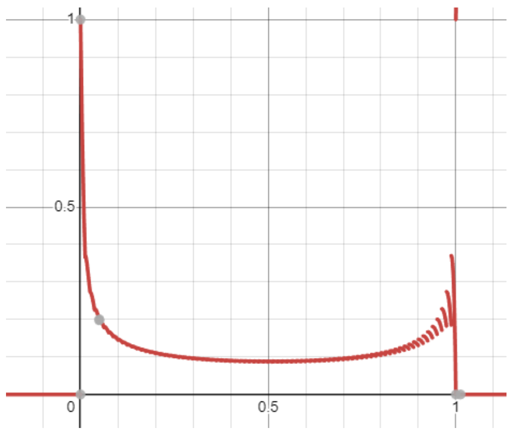


Figura 4.6: Gráfico de $f(p) = \mathbb{P}\left(\left|\frac{\sum_{i=1}^{N} X_i}{N} - p\right| \leq 0,005\right)$ para $N = 100$

De hecho, podría discutirse la situación para N mayor a 100, de modo que el rango incluya ahora a nuevos valores. El gráfico correspondiente a para la probabilidad pedida, con $N = 160$, es: Por lo cual,

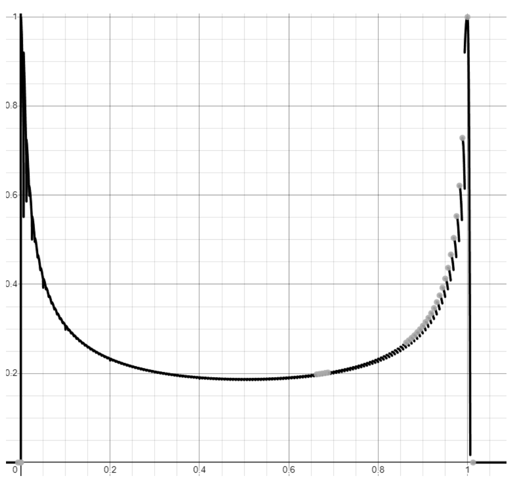


Figura 4.7: Gráfico de $f(p) = \mathbb{P}\left(\left|\frac{\sum_{i=1}^{N} X_i}{N} - p\right| \leq 0,005\right)$ para $N = 160$

quedará habilitada una cota para la probabilidad pedida, digamos por 0,198 para $N = 160$, cualquiera sea el valor de p.

Para quienes hayan utilizado la desigualdad de Bienaymé-Tchebychev, si su valor de N fue menor a 1000, esta no aportará información relevante. Será necesario que le docente proponga en discusión algún motivo para que "falle" esta técnica, con la intención de que relacionarlo con el tamaño del N.
Una vez realizada esta discusión, le docente podrá preguntar: si la probabilidad pedida es mayor a 0.2, ¿confiarían en el valor de la estimación de la proporción de bolillas blancas obtenido? La intención será que se busque un consenso respecto de qué cota inferior (nivel de confianza) les estudiantes considerarán, en ese momento, suficiente para que la estimación sea la elegida. Anticipamos que el consenso arribado podrá tomar en consideración valores para esta suficiencia que estén entre $\frac{1}{2}$ y 0,99, habilitando la propuesta del ítem c).

c) Dar un valor de N que permita confiar en la estimación realizada mediante el experimento. Luego, realizar el experimento mediante un simulador.

Aquí, el uso de la desigualdad de Bienaymé-Tchebychev permitirá a les estudiantes proponer un valor de N.
En la puesta en común de este ítem, será necesario que le docente instale en el aula la pregunta por si para cualquier nivel de confianza determinado, existirá un N que permita garantizarlo. Consideramos que los elementos de trabajo en este problema permitirán afirmar que sí, pudiendo quedar expresada en el pizarrón una escritura del tipo:
Para cualquier valor $0 < \delta < 1$, existe un N_0 tal que, para todo $N > N_0$, se tiene que

$$\mathbb{P}\left(|\frac{\sum_{i=1}^{N} X_i}{N} - p| \leq 0,005 \right) > 1 - \delta$$

El trabajo en el aula con estos asuntos posibilitará que le docente pueda definir en el aula la noción de convergencia en probabilidad, al menos en el caso particular de convergencia a una constante, ya que a lo largo del trabajo con este problema, la experiencia de trabajo de les estudiantes habrá sido en torno a eventos del tipo $P(|X_n - K| \leq \alpha)$, por lo que quedará del lado de le docente proponer la definición más general.

Conjeturamos que les estudiantes podrán arribar, posiblemente en interacción con le docente, al enunciado de la Ley Débil de los Grandes Números, al menos en una formulación general, apoyándose en lo trabajado en este problema. Consideramos interesante que le docente aborde la tarea de demostración de esta ley sin haber explicitado hipótesis particulares sobre las variables involucradas, preguntando a les estudiantes cómo procederían para demostrarla. Nos parece importante señalar que, si bien consideramos que es posible que algunes estudiantes se apoyen en lo realizado en los primeros tres ítems del problema para proponer acotar $\mathbb{P}\left(|\frac{\sum_{i=1}^{N} X_i}{N} - E(X_1)| < \alpha \right)$ mediante la desigualdad de Tchebychev, no necesariamente será algo que todes propongan. No obstante, creemos que aún en este caso y en la interacción colectiva, para estes estudiantes el uso de esta desigualdad en este contexto podrá apoyarse en lo realizado en los ítems anteriores. Vemos importante el rol de le docente en articular los pasos de la demostración colectiva con la experiencia de trabajo. Creemos, asimismo, que al proponer el uso de la desigualdad muches estudiantes podrían no haber reparado en las hipótesis necesarias para hacerlo, que son: que las variables tengan segundo momento finito, y que sean independientes. Nos parece un movimiento interesante que le docente incorpore entonces a la formulación de la Ley Débil estas hipótesis, ahora vistas como las necesarias, no solo para que la demostración sea correcta, sino con relación a los experimentos realizados hasta el momento en los problemas 1, 2, 3 y 4, en los que el rol de la independencia de las variables pudo haber sido soslayada. Esta intención de referencia con la experiencia de trabajo puede también, una vez terminada la demostración, generar condiciones para que surja la pregunta por la validez de la Ley Débil sólo con primer momento finito.

▲

4.4.5. Problema 5

Este siguiente problema se propone abordar la Ley de los Grandes Números y el Teorema Central del Límite apoyándose en algunas experiencias de trabajo que la directora de esta tesis ha tenido oportunidad de diseñar para sus aulas.
En esta propuesta, el trabajo de les estudiantes está orientado a la producción de conjeturas matemáticas apoyadas en la simulación de variables aleatorias mediante el software R. Omitimos, por cuestiones de extensión y alcance de este trabajo, discutir la especificidad de los procesos de enseñanza-aprendizaje que involucran el desarrollo de competencias específicas de programación, aunque estas sean de bajo nivel de complejidad. Asumiremos, en lo que sigue, que las consignas que deban ser realizadas en la computadora serán técnicamente posibles para les estudiantes.
En este problema, les estudiantes ya deberán haber estudiado la noción de variable aleatoria continua, y conocer las distribuciones famosas uniforme, exponencial y normal. Trabajaremos la propiedad de que suma de normales independientes es normal de los parámetros correspondientes. También asumimos que ya han trabajado con histogramas, explorando de que manera estos se relacionan con la función de densidad.

La propuesta consiste en visualizar la distribución emírica (haciendo histogramas) del promedio utilizando diferentes distribuciones para generar los datos, procuranto que esta exploración resulte un disparador para conjeturar los mencionados resultados. En una primera instancia, al generar datos bajo la distribución normal, hay un trabajo de constatación, siendo que propiedades analíticas (calculando con convolución la densidad de la suma de variables independientes) han permitido caracterizar analíticamente la distribución del promedio. Es decir, damos por sabido el siguiente hecho.

Proposición 4.4.1 *Si X, Y son dos variables aleatorias independientes que se distribuyen respectivamente con $N(\mu_X, \sigma_X^2)$, $N(\mu_Y, \sigma_Y^2)$, entonces*

$$X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2).$$

En particular, si $X_i \sim N(\mu, \sigma^2)$, $i = 1\ldots, n$, son variables aleatorias iid, entonces

(4.2)

$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

Presentamos ahora la lista de enunciados para trabajar en clase.

a) Sean $\{X_i\}_{i\in N}$ variables aleatorias iid, $X_i \sim N(\mu, \sigma^2)$, con $\mu = 70$ $\sigma^2 = 3$. Utilizaremos X de manera genérica para denotar una variable con misma distribución: $X \sim X_i$.

1. Generar $Nrep$ valores correspondientes a $\overline{X}_n$ y realizar un histograma. Notar que, para $n = 1$, $\overline{X}_n \sim X$.

2. Reperir utilizando ahora promedios calculados con $n = 1, 2, 5, 30$ y 100 valores.

3. ¿Qué puede decir sobre la distribución de $\overline{X}_n$?

Despues de que cada estudiante realice este análisis, se propone una puesta en común en la que se concluye que la distribución empírica del promedio se concentra alrededor de $\mu = 70$ a medida de n aumenta. Este hecho tiene un correlato analítico, siendo que $\overline{X}_n \sim N(\mu, \sigma^2)$. Este hecho permite calcular la probabilidad de que $\overline{X}_n$ diste de μ más que ϵ, como mostramos en la ecuación **??**. Este hecho permite intruducir la convergencia en probabilidad. De esta forma, queda enunciado una primera versión de la ley de los grandes números para variables aleatorias con distribucion normal.

Ahora bien, ¿qué pasa si repetimos el item a) generando datos con otras distribuciones?

b) Repetir el item a) generando ahora variables con diferentes distribuciones. Más específicamente

1. Considerar $X_i \sim U(67, 73)$. Calcular la esperanza y la varianza de $\overline{X}_n$.

2. Considerar $X_i \sim E(1/10)$. Calcular la esperanza y la varianza de $\overline{X}_n$.

3. Considerar $X_i \sim B(1, 0{,}7)$. Calcular la esperanza y la varianza de $\overline{X}_n$.

Esta simulación permite descubrir dos hechos: por un lado, el promedio se concentra en torno a la esperanza, y la distribución empírica parece acampanarse.

En este caso, la concentración del promedio tiene un correlato analítico en términos de la varianza del mismo, como discutimos en el Capítulo 3. Sin embargo, en estos escenarios queremos demostrar la Ley de los Grandes Números sin calcular probabilidades. En partícular, porque no se estudia la distribución del promedio de variables uniformes. En este contexto, es oportuno introducir la desigualdad de Bienaymé-Tchebychev como una herramienta útil para sortear esta dificultad, y demostrar LDGN para el caso i.i.d con segundo momento finito, lo cual permitirá dejar matemáticamente zanjada la cuestión de la concentración. Esperamos que en este momento esta desigualdad cobre otro espesor para el trabajo matemático de les estudiantes.

Habiendo dado por finalizada la discusión en torno a la concentración, vamos ahora a retomar la discusión sobre el *acampanamiento*. Nos interesará apoyarnos en el resultado, conocido por les estudiantes de que promedio de variables normales es normal, más específicamente, lo que está expresado en 4.2. Y por consiguiente,

$$\frac{\overline{X}_n - \mu}{\frac{\sigma^2}{\sqrt{n}}} \sim N(0,1)$$

Ahora bien: en el caso general, no conocemos la distribución de $\overline{X}_n$. Sin embargo, sabemos que $\mathbb{E}\left(\overline{X_n}\right) = \mu$ y que $Var\left(\overline{X}_n\right) = \frac{\sigma^2}{n}$, por lo que vamos a estudiar empíricamente los promedios estandarizados.

c) Realizar los histogramas de $\frac{\overline{X}_n - \mu}{\frac{\sigma^2}{\sqrt{n}}}$, para las diferentes distribuciones estudiadas en el ítem b). ¿Qué observa?

A partir de este trabajo, esperamos que le docente pueda introducir el enunciado del TCL, para luego abordar la tarea de demostrarlo. Consideramos relevante que le docente pueda proponer a les estudiantes la tarea (ya sea para realizar fuera del espacio de clase, o durante) de
d) Visitar el sitio Point of Significance, una publicacion de Nature dedicada a la divulgacion de la estadística dentro de las ciencias naturales. En particular, les invitamos a que miren el trabajo Importance of being uncertain, y lo relacionen con lo trabajado hasta este momento
▲

4.4.6. En torno al abordaje de variables aleatorias continuas y absolutamente continuas

En consonancia con lo propuesto en el capítulo 3, queremos proponer algunos posibles lineamientos para la entrada y trabajo en torno a este tipo de variables, que consideramos que pueden entrar en diálogo con lo que esta propuesta didáctica propone para el aula. Esto podría realizarse luego del problema 3 (donde la experiencia de trabajo en torno a la modelización probabilística de fenómenos aleatorios ya habrá tenido un despliegue que permita sostener otras experiencias de trabajo), o bien luego del problema 4. Como anticipamos, no seremos exhaustivos ni completamente detallados en este asunto, aunque esperamos que las referencias que dejaremos en las siguientes líneas, visto en el contexto de la producción de esta tesis, pueda dar herramientas para que el abordaje del diseño de la propuesta para su aula.

1. Construir experiencias de trabajo a partir de problemas modelizables mediante variables aleatorias uniformes:
   a) A partir de la idea de que la probabilidad de caer en una región solo depende del área de esa región. Una posible vía de entrada estocástica a este asunto es el conocido problema de Buffon, para el cálculo de la probabilidad de que una moneda, arrojada al azar sobre un suelo cuadriculado, toque o no los bordes de la cuadrícula[8]. En esta dirección se encuentra

[8] https://www.randomservices.org/random/buffon/Buffon.html

también el problema de la Aguja de Buffon[9].

b) A partir de problemas de modelización estocástica, como por ejemplo el de la paradoja de Bertrand, mencionado en la segunda sección de este capítulo.

2. Apelar a problemas que involucren la conceptualización de variables aleatorias continuas a partir del límite de variables aleatorias discretas. Un ejemplo de un enunciado que avanza en esta dirección es:

   Supongamos que tenemos una lámpara encendida en una habitación, y que entramos a ella cada ϵ segundos. Consideramos el experimento "entrar y que la lámpara esté apagada" y la variable aleatoria $Y =$ tiempo esperado hasta encontrarla apagada

   a) Calcular la probabilidad de que haya que esperar más de $k\epsilon$ antes de que esté apagada. Calcular la esperanza de Y.

   b) Calcular $\lim_{\epsilon \to 0} P(Y > t)$

   c) Supongamos que

   $$\frac{\epsilon}{p_\epsilon} \to_{\epsilon \to 0} \lambda$$

   Calcular la distribución de Y.

3. Discutir la noción de función de densidad, así como de esperanza y varianza, con una apoyatura que apele a la discretización de variables aleatorias (parte de lo discutido en el capítulo 3 puede ser un insumo para indagar en esta posibilidad).

4. Construir sentidos para la distribución normal. Con relación a este punto en particular, creemos que lo desplegado en el capítulo 2 en torno a la deducción de la distribución normal a partir de un razonamiento basado en la función de verosimilitud, puede ser un posible insumo. La tesis de Liliana Tauber (2001, [55]) puede ser una referencia importante para considerar.

4.5. Perspectivas y preguntas

Como se ha mencionado, sólo versiones preliminares de algunos problemas de la propuesta han sido llevados al aula. Sería interesante contar con un registro y estudio de las interacciones en un aula donde la misma sea implementada, a fin de profundizar el análisis realizado, discutir premisas y conclusiones, y elaborar en función de ello posibles modificaciones.

Avanzando un poco más en esta cuestión, nos interesa observar que la noción de convergencia en distribución puede construirse con referencia a los problemas de construcción de variables aleatorias famosas como límite en distribución de variables conocidas (como es el caso del mencionado ejemplo de las lamparitas).
A nivel contenido matemático, por su parte, la propuesta no explora problemas donde se aborden generalizaciones de la Ley de los Grandes Números o del Teorema Central del Límite, y más en particular

[9]Una posible referencia se encuentra en el artículo de Luis Santaló presente en este enlace: https://www.cienciahoy.org.ar/ch/hoy02/seccionesindiscretas.html

deja fuera de ella un espacio de reflexión en torno a la necesariedad de las hipótesis de equidistribución, finitud de momentos e independencia. Si bien es cierto que, a partir del trabajo en el aula con los problemas, se considera que quedan sentadas buenas bases para discutir en torno a estas hipótesis, cierto es que las mismas no se ven tensionadas por el trabajo con la propuesta. Asimismo, con relación a la noción de Modelización Estocástica, poder trabajar con versiones más generales de los problemas sería interesante a partir de poder ampliar el universo de situaciones modelizables. En este sentido, una investigación en torno al desarrollo de Poisson al respecto de la Ley de los Grandes Números se considera una vía fértil para, siguiendo la línea de esta investigación, producir conocimiento matemático-didáctico que permita abordar estos asuntos.

Mención aparte requiere la Ley Fuerte de los Grandes Números, que queda por fuera del alcance de esta propuesta. Esto se corresponde con una dificultad amplia encontrada durante la elaboración de esta tesis, a propósito del tratamiento didáctico-matemático de los llamados “eventos de cola”, de los cuáles uno en particular es el correspondiente a la convergencia casi segura. La complejidad matemática de su construcción, particularmente asociada al trabajo técnico sobre sigma-álgebras que requiere, y en particular la complejidad semántica involucrada en su definición y en las tareas que implican su utilización, redundaron en dejar dicho asunto para una investigación de más largo aliento. Un estudio similar al llevado adelante en esta tesis, tomando como referencia los trabajos de Borel y Cantelli, podría ser una buena vía de entrada.

Capítulo 5

Conclusiones

En la introducción de esta tesis referimos la intención primitiva de diseñar y analizar una propuesta didáctica para el abordaje de la Ley de los Grandes Números y el Teorema Central del Límite. Decidimos, en consonancia con los aportes didácticos señalados al principio del capítulo 4, realizar una investigación en torno a la historia del desarrollo de estos conceptos. Esto nos proveyó de una cantidad de elementos que fueron necesarios tanto para el análisis bibliográfico desplegado en el capítulo 3, como para el diseño de los problemas de esta propuesta. Desde nuestro punto de vista, esto da cuenta de la importancia de explorar el desarrollo de los conceptos matemáticos en la comunidad académica como un recurso valioso para el diseño de propuestas didácticas. El intento de comprender cómo estos conceptos han evolucionado a lo largo del tiempo -como vimos, siempre en relación con otros conceptos en desarrollo- proporciona elementos de interés para enriquecer las acciones sobre los procesos de enseñanza-prendizaje. En particular, la interconexión entre distintos aspectos de la matemática permite comprender otros costados de las relaciones -matemáticas y no matemáticas- y los fundamentos que subyacen en cada concepto específico.

En el capítulo 4 hemos abordado cómo la construcción de una Teoría de Probabilidades requiere que los estudiantes realicen movimientos de abstracción que pueden referenciarse en experiencias de trabajo matemático concreto. Algunos elementos a considerar en esta dirección se han esbozado en el capítulo 3, donde hemos señalado asuntos de carácter más bien general en torno a posibles organizaciones y aproximaciones teóricas a determinados conceptos y propiedades que circulan en las aulas donde se aprende y se enseña probabilidad y estadística. Consideramos fundamental proporcionar contextos y situaciones desafiantes que permitan a los estudiantes explorar y construir conocimiento colectiva e inclusivamente.

La propuesta que hemos diseñado partió de poner en un lugar central una mirada articulada entre la probabilidad y estadística, entendiéndolas como disciplinas modernamente separadas, pero con una profunda historia común que las entrelaza. En consonancia con una vasta bibliografía en torno a la enseñanza y aprendizaje de la probabilidad y la estadística, enfocamos esta relación bajo el signo de la noción de ”Estocástica", y propusimos una acepción de Modelización Estocástica compatible con otras investigaciones actuales. Este concepto nos resultó fértil para enfocar el diseño y análisis de la propuesta, apoyada en la propuesta de problemas propios de la inferencia estadística para sostener y promover la construcción de conocimiento, y en particular abordar la compleja tarea de desarrollar una Teoría de Probabilidades matemáticamente sólida.

Una perspectiva de profundización de este trabajo puede ser indagar en el rol que los argumentos que hoy la literatura denomina como "de tipo bayesiano” en el desarrollo de la probabilidad y la estadística. Nuestra investigación encontró diversas referencias a esos asuntos. Históricamente se han encontrado trazas significativas que relacionan el desarrollo de la TP con razonamientos de esta naturaleza -parte de la deducción de la curva normal realizada por Gauss involucra argumentos de este tipo-, y sin duda existen considerables capas de sentido que pueden ser desarrolladas a partir de una propuesta de investigación en esta dirección. El libro de Hald ([41]) aborda algunas de estas cuestiones. Consideramos que esta indagación puede aportar elementos para profundizar y ampliar el trabajo ralizado en esta tesis.

Esta propuesta, por su parte, contempla de parte de le docente una posición particularmente activa, incluso en las fases adidácticas, cuando se abstiene de proveer los conocimientos que les estudiantes deben movilizar para resolver determinadas situaciones. Reafirmamos la centralidad de la intencionalidad docente, tanto como la del diseño cuidadoso de los problemas.

Sin embargo, para enriquecer y problematizar estas propuestas, es necesario llevarlas a la práctica en el aula y analizar las experiencias resultantes, para lo cual el marco de la Ingeniería Didáctica referido en el capítulo anterior nos parece pertinente. El contraste con la experiencia de las situaciones de enseñanza-aprendizaje permitirá afinar y discutir las consideraciones sobre la propuesta en sí misma. Hemos esbozado en el capítulo 4 marcos de la Didáctica de la Matemática pertinentes para avanzar también en esta dirección. También hemos dejado abierta la posibilidad a considerar variaciones de esta propuesta, con la intención de que llevadas al aula puedan aportar nuevas experiencias para les estudiantes y docentes. Tenemos la esperanza de que estos nuevos insumos puedan aportar, en el complejo y fascinante mundo de la reflexión didáctico-matemática, elementos para enriquecer, profundizar, ampliar y discutir con los productos de este trabajo.

Bibliografía

[1] Artigue, M (2018). Epistemología y didáctica. El Cálculo y su Enseñanza, 11, pp. 1-31 .

[2] Artigue, M. (2000). Didactic engineering and the complexity of learning processes in classroom situations. Communication to MADIF2. Gothembourg.

[3] Artigue, Michèle; (1988). "Ingeniería Didáctica" en Artigue M, Douady R, Moreno L, Gómez P. (editor) (1995), Ingeniería didáctica en Educación Matemática: Un esquema para la investigación y a innovación en la enseñanza y aprendizaje de las matemáticas. Una Empresa Docente; Grupo Editorial Iberoamérica; Bogotá. pp. 33-60

[4] Batanero, C (2005) Significados de la probabilidad en la educación secundaria Revista Latinoamericana de Investigación en Matemática Educativa, vol. 8, núm. 3, noviembre, 2005, pp. 247-263

[5] Batanero, C; Borocvnik, M. (2016) Statistics and Probability in High School. Rotterdam/Boston: Sense Publishers, 2016

[6] Godino, J. D., Batanero, C., y Cañizares, M. J. (1987). Azar y probabilidad. Madrid,. España: Síntesis. Hacking, I. (1975)

[7] Bernoulli, Jacques (1987). Ars conjectandi-4ème partie (N. Meunier, Trans.) Rouen: IREM [Trabajo original publicado en 1713]

[8] Bertrand, Joseph (1889), Çalcul des probabilités", Gauthier-Villars, p. 5-6..

[9] Bessot, A. (2011). L'ingénierie didactique au coeur de la théorie des situations. Dans En amont et en aval des ingénieries didactiques. Margolinas, C. et al. (eds.) La pensée sauvage, Grenoble, 29–56.

[10] Billingsley, P. (1995) Probability and Measure. 3rd Edition, John Wiley and Sons, New York.

[11] Bishop, C. M., Nasrabadi, N. M. (2006). Pattern recognition and machine learning (Vol. 4, No. 4, p. 738). New York: springer.

[12] Bloch, I. (1999). L'articulation du travail mathématique du professeur et de l'élève dans l'enseignement de l'analyse en première scientifique. Détermination d'un milieu – Connaissances et savoirs. Recherches En Didactique Des Mathématiques, 19(2), 135–194. https://revue-rdm.com/1999/l-articulation-du-travail/

[13] Borovcnik, M. (2016). Probabilistic Thinking and Probability Literacy in the Context of Risk (Pensamento probabilística e alfabetização em probabilidade no contexto do risco). Educação Matemática Pesquisa. vol. 18. 1491-1516.

[14] Brousseau, G., Warfield, V. (2020). Didactic Situations in Mathematics Education. In: Lerman, S. (eds) Encyclopedia of Mathematics Education. Springer, Cham.

[15] Brousseau, G., Brousseau, N., Warfield, V (2002) An experiment on the teaching of statistics and probability. En Journal of Mathematical Behavior 20 (2002) P. 363–411.

[16] Brousseau G., Centeno J. (1991), Rôle de la mémoire didactique de l'enseignant, Recherches en didactique des mathématiques, vol.11/2.3, La Pensée Sauvage, Grenoble.

[17] Brousseau, G (1988); Los diferentes roles del maestro. Publicado en Parra,C y Saiz,I (comps) Didáctica de la Matemática. Aportes y Reflexiones. Buenos Aires, Paidós Educador, 1994.

[18] Bienaymé, I. (1867) Considérations à láppui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés. Journal de Mathématiques Pures et Appliquées, Série 2, Tome 12, P. 158-176

[19] Chin, C.W. (2021). A Short and Elementary Proof of the Central Limit Theorem by Individual Swapping. The American Mathematical Monthly, 129, 374 - 380.

[20] Cantelli F. P. (1917). La tendanza ad un limite nel senzo del calcolo delle probabilità. Rendiconti del Circolo Matematico di Palermo. 16. P. 191-201.

[21] Chevallard, Y. (1985). La transposición didáctica. Del saber sabio al saber enseñado, Buenos Aires: Aiqué, 1991

[22] Chevallard, Y (1999). El análisis de las prácticas docentes en la teoría antropológica de lo didáctico. En: Recherches en Didactique des Mathématiques. Vol. 19, nº 2, pp. 221-266

[23] K. L. Chung, (2001) "A Course in Probability Theory," 3rd Edi-tion, Academic Press, Cambridge.

[24] Fernández Coronado, N.A.; García-García, J.I.; Arredondo, E.H.; Araya Naveas (2022) I.A. Epistemic Configurations and Holistic Meaning of Binomial Distribution. Mathematics, 10, 1748. https://doi.org/10.3390/math10101748

[25] D'Amore, B. (2021). Didáctica de la matemática (2nd ed.). Magisterio. Retrieved from https://www.perlego.com/book/3187252/didctica-de-la-matemtica-pdf

[26] Daston, L. (1988). The Prehistory of the Classical Interpretation of Probability: Expectation and Evidence. In Classical Probability in the Enlightenment (pp. 3?48). Princeton University Press. https://doi.org/10.2307/j.ctv1fkgd14.5

[27] Dobler, C. (2021). A short proof of Lévy's continuity theorem without using tightness.

[28] Douady, R. (1998). Nacimiento y desarrollo de la didáctica de las matemáticas en Francia: rol de los IREM. En Artigue, M., Douady, R., Moreno, L., Gómez, P. (Eds.). Ingeniería didáctica en educación matemática. Colombia. Una empresa docente

[29] Durrett, R. (1996). Probability: theory and examples. Belmont, CA: Duxbury Press. ISBN: 0-534-24318-5

[30] Duarte, B. Lezcano Duarte, P. Maciejowski, F (2021). El azar y el manejo de la información a través de la matemática. En Entre docentes I: matemática para el aula del ciclo básico de secundaria. Compilado por Mónica Agrasar; Graciela Chemello. UNIPE: Editorial Universitaria. CABA.

[31] Esteley, C. y Magallanes, N (2021 a). Modelización estocástica en la formación de profesores en matemática. En Tendencias y nuevos desafíos de la investigación en Educación Estadística en Latinoamérica: libro de ponencias de las III Jornadas Latinoamericanas de investigación en Educación Estadística / compilación de Liliana Tauber; Jesús Pinto Sosa; editado por Liliana Tauber; Jesús Pinto Sosa. - 1a ed compendiada. - Santa Fe: Universidad Nacional del Litoral, 2021. P.55 a 62.

[32] Esteley, C. y Magallanes, N (2021 b). Modelización estocástica y empoderamiento. EnTendencias y nuevos desafíos de la investigación en Educación Estadística en Latinoamérica: libro de ponencias de las III Jornadas Latinoamericanas de investigación en Educación Estadística / compilación de Liliana Tauber; Jesús Pinto Sosa; editado por Liliana Tauber; Jesús Pinto Sosa. - 1a ed compendiada. - Santa Fe: Universidad Nacional del Litoral, 2021. P.109 a 115.

[33] Fischbein, E., (1975). The Intuitive Sources of Probabilistic Thinking in Children, D. Reidel Publishing Company, USA.

[34] Fischer, Hans. (2011). A History of the Central Limit Theorem: From Classical to Modern Probability Theory. 10.1007/978-0-387-87857-7.

[35] Gal, Iddo. (2005). Towards "Probability Literacy"for all Citizens: Building Blocks and Instructional Dilemmas.

[36] Gascón, J. y Nicolás, P. (2020). Paradigmas didácticos y reforma curricular: el caso de la teoría antropológica de lo didáctico. En: Revista Educacao Matemática Pesquisa Sao Paolo. v. 22 , n. 4, pp. 423- 437

[37] Gauss, C. F. (1809). Theoria motus corporum coelestium in sectionibus conicis solem ambientium. Perthes et Besser, Hamburg.Werke, 7 , 1-280.

[38] Georgii, H. (2013). Stochastics: Introduction to Probability and Statistics. Berlin, Boston: De Gruyter.

[39] González-Martín, Alejandro. (2005). La generalización de la integral definida desde las perspectivas numérica, gráfica y simbólica utilizando entornos informáticos: problemas de enseñanza y aprendizaje. Tesis doctoral. Repositorio institucional de la Universidad de La Laguna.

[40] Godino, J.D. (2012). Origen y aportaciones de la perspectiva ontosemiótica de investigación en Didáctica de la Matemática. En A. Estepa, Á. Contreras, J. Deulofeu, M. C. Penalva, F. J. García y L. Ordóñez (Eds.), Investigación en Educación Matemática XVI (pp. 49 - 68). Jaén: SEIEM

[41] Hald, A (2006) .A History of Parametric Statistical Inference from Bernoulli to Fisher, 1713 to 1935", Springer, NY.

[42] Jaynes, E. T. (1973), "The Well-Posed Problem", Foundations of Physics, 3 (4): 477–493,

[43] Kolmogorov, A. N. (1956). Foundations of the theory of probability. Chelsea Publishing Co.

[44] Lindeberg, J. W (1922). .Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung."Mathematische Zeitschrift 15: 211-225. <http://eudml.org/doc/167717>.

[45] Meester, Ronald. (2008). A Natural Introduction to Probability Theory. 10.1007/978-3-7643-8724-2.

[46] Metzger, W. J. (2021) Statistical Methods in Data Analysis, Facultad de Ciencias Naturales, Universidad Católica de Nimega.

[47] Rincon, L. (2007). Curso intermedio de probabilidad. Facultad de Ciencias UNAM, Mexico DF.

[48] Sadovsky, P. (2005) "La Teoría de Situaciones Didácticas: un marco para pensar y actuar la enseñanza de la matemática" en Alagia, H., Bressan, A. y Sadovsky, P.: Reflexiones teóricas para la Educación Matemática. Buenos Aires, Libros del Zorzal.

[49] Sadovsky, S (2005) Enseñar matematica hoy. Miradas, sentidos y desafios. Buenos Aires, Libros del Zorzal

[50] Santalo, L. (1980). "Estadística y probabilidad en la formación de docentes", en Proceedings of the fourth ICME.

[51] Sadovsky, P. y Sessa, C. (2005). The adidactical interaction with the procedures of peers in the transition from arithmetic to algebra: a milieu for the emergence of new questions. Educational Studies in Mathematics, 59. 85-112.

[52] Santos, J. B y Ruiz Camuñez, J. A. (2007) El problema de los dados del caballero de Méré: soluciones publicadas en el siglo XVII. En Revista SUMA N°56

[53] Székely, G. J (1986), Paradoxes in Probability Theory and Mathematical Statistics. Budapest, Akadémiai Kiadó.

[54] Seneta, E (1992). On the History of the Strong Law of Large Numbers and Boole's Inequality. Historia Mathematica, Volume 19, Issue 1. P. 24-39

[55] Tauber, L.M. (2001). La construcción del significado de la distribución normal a partir de actividades de análisis de datos. (Tesis doctoral inédita). Universidad de Sevilla, Sevilla.

[56] Tchébychef, P.-L (1867). Des valeurs moyennes. Journal de Mathématiques Pures et Appliquées, Série 2, Tome 12, pp. 177-184

[57] Trotter, H.F (1959). An elementary proof of the central limit theorem. Arch. Math 10, 226–234

[58] Uspensky, J.V. (1937) Introduction to Mathematical Probability. McGraw-Hill, New York, 19-20.

[59] Yackel, E. y Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. Journal for Research in Mathematics Education, 27(4), 390-408.

[60] Wasserman, L. (2010). All of statistics : a concise course in statistical inference. New York: Springer. ISBN: 9781441923226 1441923225

[61] Wild, C.J., y Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. International Statistical Review, 67.