



Contextos, analogías y tareas, que exponen el propósito de los conceptos clave de probabilidad

Contexts, Analogies, and Tasks that Expose the Purpose of the Key Concepts of Probability

*Manfred Borovcnik*¹

Resumo

La probabilidad es y seguirá siendo un concepto virtual. Esta especificidad requiere meta-estrategias que van mucho más allá de la instrucción de los detalles matemáticos y piden un uso sensato de la simulación. Sugerimos enfocar los esfuerzos de enseñanza en exponer explícitamente a los estudiantes al propósito del concepto de probabilidad. El propósito muestra el carácter de probabilidad indirectamente como los pasos requeridos para resolver una tarea hacen que las propiedades parezcan naturales en el contexto. Elaboramos tareas adecuadas y animaciones interactivas, que están diseñadas para superar los obstáculos de aprendizaje. Nos centramos en tres aspectos de la probabilidad: El carácter de las afirmaciones probabilísticas, el uso transparente de la probabilidad para las decisiones bajo riesgo y las consideraciones de inferencia informal en la educación temprana de probabilidad. Un criterio esencial de la enseñanza es hasta qué punto permite a los alumnos un acceso más directo a los conceptos en un nivel intuitivo. Para el diseño de animaciones didácticas, nuestros principios se caracterizan por las siguientes ideas: Se explora un cambio dinámico en comparación con la situación inicial. Como si se tratara de ver un vídeo, se observan las diferentes etapas de la emergencia de una relación entre los conceptos investigados.

Palabras clave: Interacción de intuiciones; Propósito de los conceptos; Pensamiento probabilístico; Animaciones dinámicas.

Abstract

Probability is and will continue to be a virtual concept. This specificity requires meta strategies that go far beyond the instruction of the mathematical details and ask for a sensible use of simulation. We suggest focussing teaching efforts on explicitly exposing the students to the purpose of the concept of probability. The purpose shows the character of probability indirectly as the required steps for solving a task make the properties appear natural in the context. We elaborate suitable tasks and interactive animations, which are designed to overcome learning obstacles. We focus on three aspects of probability: The character of probabilistic statements, the transparent use of probability for decisions under risk, and informal inference considerations in the early probability education. An essential criterion of teaching is how far it allows learners a more direct access to the concepts on an intuitive level. For designing didactic animations, our principles are characterised by the following ideas: A dynamic change is explored in comparison to the initial situation. Like watching a video, one looks at the different stages of emergence of a relation between the investigated concepts.

Keywords: Interplay of intuitions; Purpose of concepts; Probabilistic thinking; Dynamic animations.

Submitido em: 22/11/2019 – **Aceito em:** 27/01/2020 – **Publicado em:** 03/03/2020

¹ Doctor in Statistics and its Didactics from the University of Klagenfurt (AAU). Professor at the Department of Statistics, Alpen-Adria Universität Klagenfurt (AAU), Klagenfurt, Austria. E-mail: manfred.borovcnik@aau.at

1 Introducción – identificar los conceptos clave

A pesar de todos los esfuerzos educativos para materializar el concepto de probabilidad mediante el uso extensivo de simulaciones con el fin de ilustrar su impacto en aplicaciones repetidas, la probabilidad es y sigue siendo un concepto *virtual*. No sólo Spiegelhalter (2014a, b) usó la redacción de probabilidad como metáfora; es decir, necesitamos imágenes y concepciones visuales cualitativas en nuestra mente para hablar de probabilidad y entender su significado. Un indicador de las dificultades con probabilidad también puede verse en la abundancia de conceptos erróneos (por ejemplo, Kahneman, Slovic & Tversky, 1982) y paradojas (Székely, 1986).

Hay varios estudios y monografías que revelan el complejo campo de la probabilidad visto desde una perspectiva más amplia de la filosofía; por nombrar sólo algunos aquí: Batanero, Henry y Parzysz (2005), Batanero, Chernoff, Engel, Lee y Sánchez (2016) y Bennett (1999). Sobre la diversidad de las respuestas educativas, el lector puede orientarse por Chernoff y Sriraman (2014), centrándose desde el tema general en el *pensamiento probabilístico*, y Jones, Langrall y Mooney (2007), con investigaciones sobre la probabilidad en comparación con las realidades en el aula, es decir, los abundantes problemas conocidos en la enseñanza de la probabilidad.

Las intenciones didácticas para explicar o aclarar la naturaleza de la probabilidad son múltiples. Steinbring (1991) desarrolla la idea de probabilidad como un concepto teórico que se desmorona y se queda sin un significado más profundo si sólo se enseñan aspectos específicos dentro de un enfoque de enseñanza. Steinbring se refirió explícitamente a las implicaciones negativas de centrarse demasiado estrechamente en la equiprobabilidad de Laplace y descuidar el aspecto de las frecuencias relativas y viceversa. Carranza y Kuzniak (2008) hablan de la dualidad de la probabilidad y se refieren a los aspectos inseparables de las frecuencias relativas y a las ideas subjetivistas de probabilidad; trazan las graves desventajas de descuidar los aspectos subjetivistas de la probabilidad.

Podemos utilizar el término complementariedad en lugar de dualidad para destacar la conexión inseparable de un concepto con sus diversas partes. El término se origina en la discusión de la teoría cuántica en la década de 1920; ha sido introducido por Niels Bohr, quien declaró que el conocimiento completo de los fenómenos en las dimensiones atómicas requiere una descripción de las propiedades tanto de las ondas como de las partículas y que no puede reducirse a un solo aspecto sin que se pierda completamente el significado (Bohr, 1934/1927). La complementariedad se ha transferido a la didáctica de las matemáticas en los años de 1980 para describir el dilema didáctico resultante de los aspectos inseparables de los conceptos (véase Borovcnik, 1992).

Otra complementariedad de la probabilidad puede verse en sus estrechas interrelaciones con la inferencia estadística. A finales de los años ochenta se produjo una revolución en la didáctica de la estocástica, en la medida en que fue testigo de los primeros enfoques serios hacia la inferencia estadística, cuando los investigadores reconocieron la

fuerte conexión entre probabilidad e inferencia estadística (Borovcnik 1992): las estadísticas descriptivas ya no se consideraban suficientes para aclarar cuál debía ser el papel de la probabilidad en la enseñanza; por el contrario, el significado de los conceptos probabilísticos se amplía y aclara dentro de un marco inferencial: La probabilidad sin métodos de inferencia estadística es nula, mientras que la inferencia estadística no puede ser entendida sin probabilidad. El propósito de la probabilidad – en este sentido – siempre ha sido diseñar métodos de inferencia estadística que permitan generalizar el conocimiento parcial (sobre una “muestra”) a una entidad más amplia (llamada población, que podría describirse mediante una distribución de probabilidad).

Gigerenzer (2002) es un hito para vincular la probabilidad con el riesgo, que ha sido un doble compañero de la probabilidad desde sus primeros desarrollos. La relación entre probabilidad y riesgo es tan estrecha que, tras una inspección más detallada, no se podrían separar sus aspectos, lo que indica una vez más una complementariedad. Borovcnik (2016) llega incluso a rediseñar la alfabetización probabilística a la luz del riesgo. Reformular el propósito de la probabilidad como encontrar o inventar métodos para encontrar decisiones razonables bajo riesgo mejora mucho la percepción de probabilidad.

Con el fin de superar el impacto negativo de los enfoques unilaterales de la enseñanza de la probabilidad y la estadística (y de otras disciplinas de la ciencia y las matemáticas), Fischbein desarrolló sus ideas sobre las intuiciones primarias y secundarias (Fischbein, 1975, 1987). Para el uso didáctico de desarrollar una interacción entre las intuiciones primarias crudas que personas tienen antes de cualquier enseñanza y la exposición a tareas matemáticas que requieren la aplicación de los conceptos que los estudiantes aún no han desarrollado, se puede volver a la obra de Fischbein o a Borovcnik (1992).

En Borovcnik (2019b) se ha establecido un marco teórico para identificar las ideas fundamentales con probabilidad y desarrollar intuiciones sostenibles. Se ha defendido el papel eminente de la interacción entre las intuiciones primarias en bruto y las intuiciones secundarias emergentes que los estudiantes adquieren por su trabajo continuo con las tareas. Las intuiciones secundarias son el objetivo final de los esfuerzos de enseñanza: este tipo de intuiciones no sólo neutralizan las intuiciones primarias engañosas (que de otro modo serían muy resistentes al cambio mediante intervenciones de enseñanza), sino que también apoyan una comprensión holística (“ganzheitlich”) de las soluciones a través de los métodos matemáticos. Un entendimiento, que funcionaría también sin conocer cada detalle matemático de la forma en que se resolvió la tarea. Un vínculo entre los conceptos matemáticos y las intuiciones secundarias puede establecerse por analogías (sobre el papel de las analogías, se puede consultar Simons, 1984).

Hacer que el propósito de los conceptos a ser aprendidos también introduciría un impulso en la comprensión y aceptación de los nuevos conceptos y los métodos que se basan en ellos. El propósito de un concepto es enfocar inmediatamente los pasos de aprendizaje requeridos; hace que los pasos sean comprensibles y los hace “naturales”. Las matemáticas que hay que aprender adquieren el carácter de una herramienta y uno se familiariza más con

la herramienta, cuanto más la utiliza. Trabajar con el concepto como una herramienta para servir a un propósito también evita la discusión interminable de cómo se puede entender un concepto (lo cual puede ser realmente confuso, especialmente para los alumnos en su etapa inicial) al mostrar para qué se puede usar el concepto. Si el propósito les parece razonable a los alumnos, estarán más dispuestos a seguir los pasos de aprendizaje que se requieren para completar su red cognitiva para su comprensión conceptual de los conceptos que ya han sido utilizados por ellos ampliamente.

Esto sigue también más de cerca la forma en que los conceptos han sido utilizados en la fase de su aparición: han sido utilizados implícitamente para un propósito específico. Es una lástima que hayamos perdido la conexión con el origen de la probabilidad y su conceptualización dentro de la física, ya que la probabilidad sirve genuinamente para resolver el objetivo de comprender los fenómenos físicos estructurándolos mediante teorías emergentes, y la probabilidad ha sido una piedra angular de la teorización de la física (véase, por ejemplo, von Plato, 1994).

Revivir ese alto nivel de los conceptos, para centrarse en el propósito de los mismos, puede ser visto como un objetivo primordial de las intervenciones de enseñanza. Utilizamos una forma hermenéutica de argumentar, y relacionamos las ideas fundamentales de probabilidad con sus antecedentes matemáticos y filosóficos; el objetivo de este trabajo es hacer transparente el propósito de la probabilidad mediante el desarrollo de tareas que sean adecuadas para ese “propósito”. Como la probabilidad es un concepto teórico, introducimos la idea del propósito del concepto y de varias tareas que pueden ser manejadas con el concepto. Tal vez sea mejor dejar claro con qué propósito se debe utilizar el concepto en lugar de explicar y explicar con más detalle qué propiedades tiene el concepto. El objetivo final de las intervenciones de enseñanza debería ser facilitar *un acceso intuitivo a los conceptos clave* para los alumnos.

- Explorar el tipo y la calidad de la información de probabilidad (capítulo 2).
- Propósito de la probabilidad – comparar riesgos y preparar decisiones (capítulo 3).
- Explorar un modelo probabilístico para tomar mejores decisiones (capítulo 3).
- Medir o estimar una cantidad desconocida (capítulo 4).
- Aprender de la teoría – teoremas centrales.
- Actualización de un juicio cualitativo de probabilidad por datos (enfoque bayesiano).

El primer punto toca el núcleo del carácter de las declaraciones de probabilidad sobre las que hay tanta confusión. Los puntos segundo y tercero (y el último también) se refieren a la toma de decisiones y a un componente cualitativo de la probabilidad y se refiere a la obtención y actualización de información cualitativa sobre la credibilidad de las declaraciones o eventos. El punto cuarto establece una conexión temprana del concepto de probabilidad con la inferencia estadística, que es uno de los principales propósitos de la

probabilidad: justificar las decisiones entre opciones modeladas por la probabilidad. El quinto punto prepara una conexión entre intuiciones primarias e intuiciones secundarias (Fischbein, 1975, 1987) y proporciona una estructura adecuada para las intuiciones que deben surgir de la enseñanza de partes de la matemática de la probabilidad y la estadística. También debe fortalecer la conexión entre la probabilidad y la inferencia estadística, que a su vez revisa las concepciones de probabilidad en un “último paso educativo”.

Los dos últimos puntos de la lista anterior no se tratarán más adelante en el presente documento. El proyecto “Conceptos fundamentales y sus propiedades clave en probabilidad” tiene otros artículos en preparación sobre estos aspectos, que también están contenidos en Batanero y Borovcnik (2016), Borovcnik (1992, 2015a), o Vancsó (2009, 2018).

A continuación, nos dirigimos a los estudiantes o aprendices de forma genérica. Sin embargo, definitivamente apuntamos a los estudiantes de los últimos años del gimnasio. Hemos experimentado con las ideas presentadas con estudiantes de estudios no matemáticos a nivel de bachillerato en la universidad (estos estudiantes no están más inclinados a aprender matemáticas ni son más talentosos o capaces que el grupo objetivo). Los contextos y las tareas que aquí se discuten son comparables a los de Batanero y Borovcnik (2016), *La estadística y la probabilidad en la escuela secundaria*. Las tareas se originan en un proyecto más amplio sobre Ideas Fundamentales en Estocástico. Un objetivo clave de este proyecto es desarrollar metaconocimientos para los alumnos que puedan sustituir en parte a las relaciones y teoremas matemáticos de la disciplina, que a veces resultan bastante complejos.

La probabilidad se define matemáticamente por medio de axiomas y el fundamento axiomático del concepto justifica su interpretación. Hay dos interpretaciones principales de la probabilidad, la frecuentista (FQT) y la subjetivista (SJT), matemáticamente puestas en una base sólida por Kolmogorov (1933) o de Finetti (1937/1992). La tercera interpretación relevante es la concepción de equiprobabilidad (Laplace, o APT, teoría a priori de la probabilidad), que ha sido una idea puntera en el desarrollo de la física (especialmente en la termodinámica) pero que no ha podido ser utilizada para una base sólida del concepto de probabilidad desde un punto de vista matemático. Ver Borovcnik y Kapadia (2014) para una discusión de estas interpretaciones de la probabilidad desde un punto de vista educativo y para la notación especial (FQT, SJT y APT).

El propósito de la probabilidad en el contexto de una tarea hace un uso extensivo de una o más de estas interpretaciones de probabilidad, que pueden coexistir incluso en la misma situación (lo que hace que la probabilidad no sea más fácil de entender). Las consideraciones sobre la calidad de la información probabilística (capítulo 2) se refieren a la equiprobabilidad de Laplace (APT); la comparación de riesgos y la búsqueda de mejores decisiones se refiere a la utilidad y a la interpretación subjetivista de la probabilidad (SJT). Las visualizaciones en las tareas sobre la medición de una probabilidad desconocida están vinculadas más explícitamente a la probabilidad como concepto de frecuentista (FQT). El trabajo que se está llevando a cabo en el marco del mencionado proyecto sobre las ideas fundamentales es una

elaboración sobre la inferencia estadística, que se divide abiertamente entre los métodos frecuentista (inferencia clásica) y subjetivista (inferencia bayesiana).

Como una observación adicional, es importante observar que cuando usamos el método de simulación, generamos frecuencias relativas a probabilidades relacionadas que – en la primera vista – vinculan las ilustraciones y la interpretación de la probabilidad obviamente a un concepto de frecuentista. Sin embargo, el método de simulación se utiliza ampliamente dentro del enfoque bayesiano de inferencia para determinar soluciones (también en forma de distribuciones de probabilidad). Esto significa que a pesar de la visión de probabilidad que se tiene (y esta visión es subjetivista en el marco Bayesiano), se puede aprovechar el método de simulación. Aplicar la simulación conlleva la posibilidad de ser malinterpretado por los alumnos, que entonces podrían asociar aspectos frecuentistas a conceptos subjetivistas. Por lo tanto, el método de simulación tiene que ser usado con cuidado para evitar tal cambio en la connotación de probabilidad a un rincón de interpretación frecuentista (y por lo tanto sesgada).

Con las tareas, que se discuten en las secciones siguientes, cubrimos todos los aspectos de la probabilidad incluyendo las diversas concepciones y tocamos los aspectos de la probabilidad que construyen el fondo de las propiedades clave del concepto de probabilidad. Como hemos dicho, nuestro objetivo es que los alumnos comprendan estas propiedades clave, ya que parecen naturales, surgiendo del contexto y el propósito de las tareas utilizadas.

2 Exploración del tipo y la calidad de la información de probabilidad

La probabilidad es un concepto teórico y un valor específico de la probabilidad de un evento bajo escrutinio es simplemente un número descriptivo, un índice que debería transmitir un *potencial*. Esta información de referencia se expresa de manera diferente en situaciones reales. Es esencial tener una sensación de cuán grande es el tamaño de la variación; en otras palabras, cuán grande es el margen de variación o error de las frecuencias relativas que son el único instrumento disponible para que podamos materializar el índice que llamamos probabilidad.

Mostramos un escenario de simulación que mejora la noción de margen de error, lo que proporciona una intuición estable sobre el efecto de la longitud de una serie aleatoria: cuanto mayor es la base de datos, menor es la variación de las frecuencias relativas. Mientras que el primer experimento se encuentra en la fase inicial de la enseñanza de la probabilidad, el segundo revela la variación de una muestra de una distribución normal. Contrariamente a las expectativas generales, las muestras de una distribución se parecen a la población en cuanto a la forma sólo si el tamaño de la muestra es muy grande.

2.1 El margen de error de los eventos aleatorios

Una tarea principal en las primeras fases de la educación en probabilidad es aclarar y apoyar la relación entre las probabilidades y las frecuencias relativas de eventos en ensayos

repetidos de un experimento aleatorio. Esta relación se justifica más tarde en el curso por la ley de grandes números, que describe una conexión sofisticada, que es completamente diferente de la convergencia en el cálculo, pero que siempre atrae ideas engañosas sobre la convergencia. Exploramos la variación de la aleatoriedad (no la convergencia) para mejorar la comprensión del tipo y la calidad de la información de probabilidad en lugar de ilustrar una ley de convergencia, lo cual no se puede hacer en ninguna serie finita. Nuestro objetivo es mostrar un patrón de reducción en la variabilidad de las frecuencias relativas. Un experimento adecuado para este propósito es seleccionar dígitos aleatorios de 0 a 9.

Una dificultad básica es entender el tipo de información que se encuentra en una declaración de probabilidad. ¿Qué significa realmente una probabilidad de $\frac{1}{2}$? Se puede vincular a una decisión justa entre dos posibilidades. Por otra parte, una situación sin preferencia se vincula demasiado fácilmente con probabilidades iguales; este sesgo de equiprobabilidad de Lecoutre (1992) parece más incitante si hay dos posibilidades. La mejor manera de aclarar esta cuestión es mediante un debate abierto en forma de entrevista empírica con situaciones que varían (véase Borovcnik & Peard 1996).

¿Cómo se relaciona esta expresión con los experimentos repetidos? Es vital *evitar un enfoque demasiado fuerte en el patrón en el que se desarrolla la serie*, ya que, si se aplica la hipótesis de la aleatoriedad, cada patrón es igualmente probable, salvo que uno considere el patrón como una clase (un grupo) de resultados. Hay mucha investigación sobre las secuencias aleatorias y el juicio de la probabilidad de las secuencias que tienen que ser comparadas (ver, por ejemplo, Chernoff 2013, o Borovcnik & Bentz 1991, 1990, 2003). Para la probabilidad como concepto, la especulación sobre los patrones es irrelevante siempre y cuando no se considere violada la hipótesis de la aleatoriedad y si es así, se trata de una prueba estadística bastante complicada de aplicar en una réplica de los datos (¡no en los mismos datos!).

En el siguiente experimento, se retoma una idea de Freudenthal (1972); Freudenthal sugirió investigar el margen de variación con un tamaño de muestra fijo en lugar de ilustrar un límite oscuro de frecuencias relativas hacia la probabilidad desconocida repitiendo el experimento aleatorio básico una y otra vez. Para una comprensión conceptual de una declaración de probabilidad es vital entender que el rango de variación se reduce con series más grandes y que, al mismo tiempo, los ensayos individuales están completamente sujetos a su carácter aleatorio. Esto significa también que no existe una regla de compensación, lo que a menudo se cree después de esfuerzos docentes para demostrar la convergencia estocástica de las frecuencias relativas hacia la probabilidad subyacente.

Simulamos dígitos aleatorios (0 a 9), cada uno con una probabilidad de $\frac{1}{10}$. En el laboratorio estadístico, podemos garantizar que esta probabilidad se mantiene para cada uno de los dígitos individuales, lo que excluye las concepciones burdas sobre los favoritos (utilizamos dígitos aleatorios de 0 a 9 y no los números cargados emocionalmente de 1 a 45 de la lotería estatal). No investigamos el desarrollo de las frecuencias relativas con la longitud de la secuencia aleatoria. Fijamos dos instantáneas en el proceso aleatorio, después de 50 y

después de 1000 dígitos son generados. Investigamos la distribución de los dos escenarios y los comparamos. Tiene que ser declarado explícitamente que normalmente no tenemos tales números (aunque hay algunas personas que leen las estadísticas de los números recientemente sorteados en la lotería estatal).

Lo que se puede ver (en la Figura 1) es que los dígitos a continuación varían de manera errática, lo que muestra que la aleatoriedad en el fondo está funcionando plenamente. Es esclarecedor repetir los escenarios de simulación completos (con 50 y 1000 dígitos) – como en un vídeo – para ser testigo de que los dígitos que están “a la cabeza” en una simulación, no están detrás en la siguiente, o viceversa. Las frecuencias de los dígitos individuales realmente varían sin ningún patrón y una vez que uno reconoce un patrón se pierde en las siguientes repeticiones del escenario de simulación.

Lo que podemos reconocer es que la variación de las frecuencias de los dígitos individuales permanece un tanto dentro de una banda (resaltada en la Figura 1) que refleja el margen – el tamaño – de la variación. Rara vez un escenario de simulación tiene frecuencias fuera de esta banda. Tenemos una impresión del margen de fluctuación aleatoria. Lo que también se puede ver es que este rango de variación es considerablemente menor con 1000 dígitos aleatorios que con sólo 50. Un patrón estable de frecuencias relativas emerge con una fluctuación aleatoria completa para cada ensayo, lo cual no es una contradicción vista desde este ángulo. También muestra que – suponiendo que la muestra se tome de hecho al azar (que es la ventaja de nuestro laboratorio estadístico, donde garantizamos que se cumplen los supuestos) – cuanto mayor sea el tamaño de la muestra, menor será la variabilidad.

Como un experimento de pensamiento, ya en este experimento inicial, uno puede especular sobre un encogimiento hacia la línea de referencia de 1/10 ya que las probabilidades para los dígitos aleatorios son las mismas desde el principio. Mostramos una simulación en la Figura 1, pero el lector debe tener en cuenta que la imagen estática sólo tiene vida si la repetición de todo el escenario de la simulación se reproduce como una película animada para demostrar el efecto que se acaba de describir.

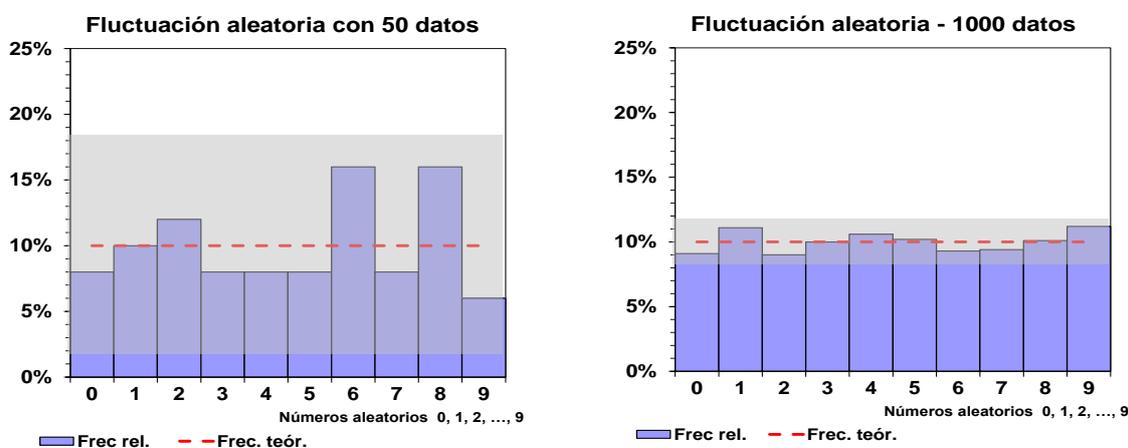


Figura 1 – Rango de variación aleatoria de dígitos aleatorios (1, 2, ..., 9) con pocos y muchos datos

Fuente: Elaborado por el autor

2.2 Muestreo de una distribución normal

Se utilizan muestras aleatorias para estimar parámetros desconocidos de la “población” (o distribución parental). Aunque este método ya funciona para tamaños de muestra relativamente pequeños, la forma de los datos generados no se parece a la distribución de la población a menos que tengamos muchos datos simulados. A continuación, investigamos dos escenarios de simulación (Figura 2):

- En el Escenario 1, generamos 50 datos de una distribución normal con una media preespecificada y una desviación estándar.
- En el Escenario 2, investigamos el mismo proceso y generamos 1000 datos.

En ambos escenarios, dibujamos un histograma para investigar la distribución empírica de los datos generados. Es mejor cambiar el valor medio o la desviación estándar para ver el efecto sobre la distribución empírica de los datos: se verá que el histograma que representa los datos generados cambiará de acuerdo con el nuevo valor de la media y mostrará menos anchura si se reduce la desviación estándar del modelo. Lo que es más importante es ser testigo de la gran variedad en la forma de la distribución empírica de los datos, que no se parece en nada a una forma normal. Incluso para el escenario con 1000 datos, la forma muestra anomalías aquí (una fuerte asimetría a la izquierda) con respecto a una distribución normal.

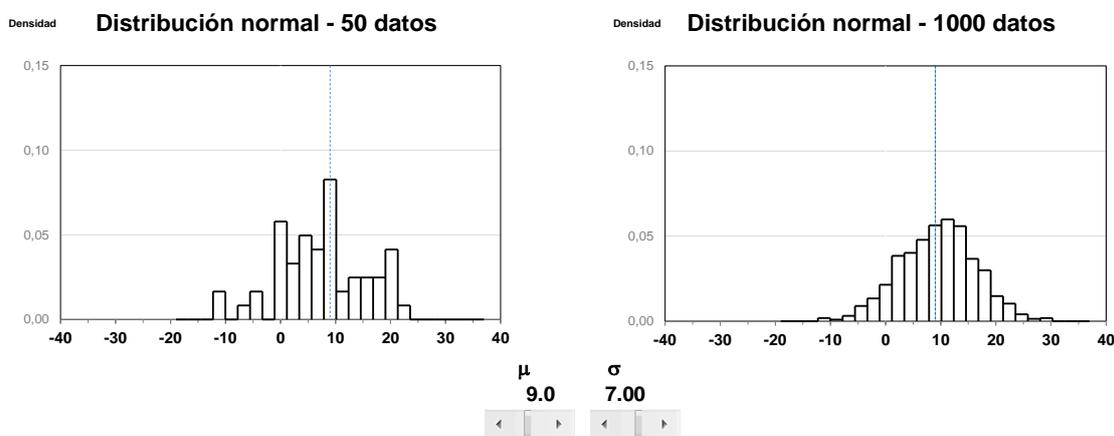


Figura 2 – Variación de muestras de una distribución normal – izquierda con 50 datos; derecha con 1000 datos

Fuente: Elaborado por el autor

Los escenarios revelan el “margen de variación” de un modelo de probabilidad. Incluso si la distribución principal se distribuye normalmente (lo que está garantizado por nuestro laboratorio estadístico), los datos están lejos de parecer como si se distribuyeran normalmente. La repetición del escenario de la simulación permite a los alumnos presenciar la variación como en un vídeo. Como resultado, vemos la relevancia del tamaño de la muestra y que con muestras pequeñas las conclusiones son inestables. Reconocemos que suposiciones como una distribución normal no pueden ser probadas estadísticamente de manera adecuada. Las variaciones que no conducen a un rechazo de la presunción de

normalidad son demasiado amplias y pueden ser causadas por una distribución parental diferente.

Sin embargo, es posible sacar conclusiones sobre varios parámetros de la población a partir de muestras más pequeñas. Como consecuencia del teorema del límite central, la distribución de varias estadísticas extraídas de una muestra es aproximadamente normal. Esto resulta ser un concepto artificial, ya que normalmente sólo tenemos una muestra de la población y, por lo tanto, nunca experimentamos la distribución del muestreo de una estadística como la media de una muestra.

3 Tareas que exponen el propósito de la probabilidad

En este capítulo se tratan las tareas que revelan el propósito de la probabilidad sin conexión explícita con la inferencia estadística. La probabilidad puede ser utilizada para hacer transparentes las decisiones bajo riesgo, para aclarar los criterios que se utilizan para definir lo que una buena decisión requiere y cómo proporcionar una solución.

Una tarea básica se refiere a un propósito primordial de la probabilidad, que es fijar el precio de la incertidumbre o el riesgo: Ya sea que juguemos a un juego de azar o contratemos una póliza de seguros, estructuralmente la situación es la misma: dos partes interesadas se reúnen e intercambian sus papeles: uno de los socios se deshace de una situación incierta y el otro asume este papel a partir del primero, es decir, está dispuesto a dejar la posición de certeza. Introduciendo un índice llamado probabilidad, que “mide” el grado de riesgo, es posible encontrar un precio para el intercambio de los roles (primera sección a continuación).

Si uno está dispuesto a referirse a la probabilidad de un resultado futuro, entonces también podría ser capaz de mejorar una decisión que tiene que ser tomada ahora. Los criterios de optimización sin incluir tal índice de probabilidad podrían llegar a conclusiones incómodas y tendrían un impacto conservador en el comportamiento: la probabilidad aumenta el alcance de las estrategias con un alto potencial de innovación pero conllevan un riesgo inherente – esta es una invariante general de la utilización de la probabilidad incluso si se puede utilizar para una ventaja global (segunda sección más adelante). Al utilizar la probabilidad como precio de la incertidumbre, hay que obedecer algunas reglas básicas, que parecen muy naturales en el contexto del riesgo (tercera sección a continuación).

Las siguientes fuentes contienen más información sobre las propiedades clave de la probabilidad: Borovcnik (2015a), Batanero y Borovcnik (2016). Cubren el aspecto de los applets dinámicos que pueden mostrar – más allá y paralelamente a las consideraciones matemáticas – las propiedades clave del concepto de probabilidad.

3.1 Fijar el precio de lo desconocido

La probabilidad sirve para intercambiar la incertidumbre y el riesgo que implica con el dinero. Esto es parte del contrato de seguro donde el cliente se enfrenta a la posibilidad (riesgo) de un accidente y paga la prima del seguro a la compañía de seguros. El cliente abandona la posición de incertidumbre (sobre las implicaciones financieras de un accidente)

DOI: 10.20396/zet.v28i0.8657607

para entrar en una posición de certeza (sin riesgo) pero paga por adelantado por ello. La compañía de seguros – por otro lado – deja la posición de certeza y asume el riesgo del cliente y recibe un pago por ello. Las probabilidades son la clave para determinar el precio del contrato. El ejemplo sirve también para discutir las diversas interpretaciones de la probabilidad (subjetivista para el cliente, frecuentista para la empresa) y la utilidad del impacto (utilidad para el cliente, dinero simple para la empresa). Para más detalles, ver Batanero y Borovcnik (2016).

Las probabilidades son la clave para determinar el precio del contrato, o la base para la toma de decisiones si el resultado depende de probabilidades, o mejor dicho, si está modelado por probabilidades. Los actores que se reúnen en una decisión (como en el contrato de seguro) no necesitan usar el mismo modelo de la situación y si usan el mismo modelo, podrían elegir diferentes probabilidades; además, si tienen los mismos valores para la probabilidad, podrían agregarle una connotación diferente: para la compañía de seguros es una probabilidad frecuente del pasado promediada sobre muchos contratos, para el tomador de la póliza es una probabilidad personal y cualitativa sobre el evento en cuestión y el tomador de la póliza no tiene un punto de vista de promediación, ya que él toma la prima sólo una vez. El “riesgo” se encuentra al cliente con toda la gama de variabilidad, mientras que la compañía de seguros soporta casi cualquier riesgo, ya que su resultado es muy estable y, por tanto, previsible debido a la promediación.

3.2 Optimizar una decisión en una situación de incertidumbre

Presentamos un ejemplo de toma de decisiones sobre el número de copias a producir para una revista si la demanda está modelada por probabilidades y hay algunos datos adicionales sobre el costo de las copias (ver Tabla 1, que se reproduce de Borovcnik 2015b); el precio por edición asciende a 1,60. Las opciones son producir 1000, 2000, ... o 5000 copias. ¿Cuál es la mejor opción?

Tabla 1 – Opciones de producción, coste de estas opciones y probabilidades de la demanda de la revista

Opción	1	2	3	4	5
Copias para producir	1000	2000	3000	4000	5000
Coste de producción	2000	2200	2400	2600	2800
Demanda	1000	2000	3000	4000	5000
Probabilidades	0,40	0,30	0,20	0,06	0,04

Fuente: Elaborado por el autor

Existen diferentes criterios para optimizar el número de copias. Se podría minimizar la pérdida máxima, por ejemplo, o se podrían optimizar los beneficios esperados. Es fácil ver (ver Borovcnik 2015b, p. 128) que la solución minimax conduce a decisiones extrañas:

“Es obvio que nadie estaría dispuesto a decidirse por la opción 1000 ya que – cualquiera que sea la demanda – llevará a una pérdida de –400. Esto nos recuerda el principio de evitar una pérdida segura [...]. La opción 2000 ofrece un beneficio esperado positivo de 360; sin embargo, también puede conducir a una pérdida aún mayor de –600 en comparación con la decisión de realizar 1000 copias. Esto refleja

DOI: 10.20396/zet.v28i0.8657607

una propiedad básica de [...] decisiones en situaciones de riesgo. Rara vez se pueden encontrar decisiones, que son mejores en todo [...]. [...] las acciones restantes [...] no pueden compararse entre sí sin un criterio adicional y cuál es la mejor decisión depende del criterio utilizado. Mejorar una situación en un aspecto (tener un mayor beneficio neto esperado) va acompañado del riesgo de mayores pérdidas potenciales. Incluso se puede hablar de una invariante en la vida humana vista desde una perspectiva filosófica general sobre el riesgo. La opción 3000, que arroja un beneficio previsto de 640, es mejor. Sin embargo, asume el riesgo de una pérdida de -800 [con una probabilidad de 0,40]. Resulta que la opción 3000 produce el máximo beneficio esperado (640) y es, en el modelo actual, la mejor decisión.”

Los conceptos implicados van desde el frecuentista hasta el subjetivista, pasando por la probabilidad y la utilidad, e incluyen una discusión sobre los criterios que se siguen cuando se quiere optimizar una decisión. Cabe señalar que la tarea de optimizar una decisión no está bien definida mientras no se fije el criterio de decisión y cada criterio tenga sus propios méritos e inconvenientes.

3.3 El valor esperado es una noción aditiva – La dispersión no siempre es aditiva

Hay varias maneras de derivar el valor esperado de una distribución binomial, que es el resultado de un experimento binario repetido n veces independientemente con una probabilidad de p para el evento en cuestión. Es muy fácil reconocer que para un juicio el valor esperado es p . Si la situación está incrustada en una apuesta – ganar 0 si el evento no ocurre y ganar 1 si el evento ocurre – el precio justo o valor esperado para esta apuesta asciende a p . Si uno hace tal apuesta n veces, entonces el precio es n veces p , lo cual es muy fácil de aceptar sin ninguna otra matemática. Sin embargo, si las apuestas son dependientes, entonces esta aditividad – aunque altamente contraintuitiva – sigue siendo válida, como se puede ver en Batanero y Borovcnik (2016).

Una relación de aditividad comparable para la varianza sólo es válida en el caso de experimentos aleatorios independientes, como lo es la suposición para los ensayos individuales que conducen a la distribución binomial. Esta aditividad se compara de alguna manera con el teorema de Pitágoras y la independencia se convierte en una relación ortogonal para las variables aleatorias. Esta es la razón que explica la preferencia por el cuadrado de la desviación estándar en favor de otras medidas de variabilidad y no cualquier consideración relativa a la calidad de la descripción de la anchura de una distribución.

Por cierto, para distribuciones de probabilidad sesgadas, ni el valor esperado ni la dispersión (o la desviación estándar) proporcionan una descripción adecuada de la ubicación o la extensión de una distribución de probabilidad. Sin embargo, estos parámetros se utilizan en todas partes. La razón de esto radica en las interrelaciones matemáticas, principalmente en el hecho de que la media de las muestras como estadística está – siguiendo el teorema del límite central – aproximadamente distribuida normalmente. Esta circunstancia simplifica los procedimientos de inferencia estadística cuando se quiere sacar conclusiones de una muestra aleatoria a la población.

4 Medir o estimar una cantidad desconocida

En esta sección, tratamos de cerrar la brecha entre la probabilidad y la inferencia estadística. Ya en las primeras fases de la enseñanza de la probabilidad, uno debería introducir conexiones informales para inferir para que el propósito de la probabilidad sea más claro. Para ello, desarrollamos una analogía entre la tarea de estimar una probabilidad (que normalmente es desconocida) con la medición repetida de una cantidad física. Tenemos que definir un procedimiento de medición y aclarar cómo podemos conceptualizar la calidad de nuestro instrumento de medición. Un mejor instrumento no siempre proporcionaría un valor de mejor medición, pero en general resultaría una ventaja si se utilizara varias veces (repetidamente). Esto significa que tenemos que introducir métodos para conceptualizar el margen de error de la medición (primera sección a continuación).

Luego, redefinimos el experimento repetido de lanzar monedas (u otro experimento binario) en el contexto de la medición. La frecuencia relativa de los cabezales en una serie de lanzamientos de la moneda se considera una medida de la probabilidad desconocida (segunda sección a continuación). En lugar de aumentar el número de lanzamientos, investigamos tres “instrumentos” por su calidad de medición, y fijamos su número de lanzamientos con 5, 10 y 20 solamente. Sólo queremos demostrar que el instrumento de medición con la serie más grande de lanzamientos puede ser considerado como más preciso (tercera sección más abajo). Y de ahí sacamos nuestra conclusión por “extrapolación” y no por un experimento material. Este experimento de pensamiento puede aplicarse a cualquier otra tarea de estimar un parámetro de una población, como la media de una población.

Bootstrap es un método moderno para estimar un parámetro desconocido de una población, que puede ser utilizado alternativamente a un intervalo de confianza. Bootstrap es uno de los métodos que se utiliza en “Inferencia Informal”, un enfoque moderno sugerido en la investigación de la estadística docente, que pretende simplificar la inferencia estadística (algunos incluso dirían que para reemplazar la inferencia estadística; ver Borovcnik, 2019 a). El contexto de medición también destaca lo que realmente se hace en una estimación de Bootstrap (cuarta sección a continuación).

4.1 Analogía entre la medición de cantidades físicas y virtuales

La probabilidad como límite de las frecuencias relativas es una percepción ingenua que también ha sido (y sigue siendo) el foco de las definiciones de probabilidad. Esto constituye la base de la vista de probabilidad del frecuentista. Es una tarea importante de la fase introductoria de la educación en probabilidad aclarar esa relación entre la probabilidad y las frecuencias relativas. El experimento habitual es repetir un ensayo aleatorio binario muy a menudo y mostrar el comportamiento limitante del desarrollo de las frecuencias relativas del evento investigado. También aquí, para evitar el fenómeno empíricamente inalcanzable de la convergencia, preferimos la variante Freudenthal (1972) del experimento, que investiga la distribución (!) del resultado del experimento a un tamaño de muestra específico cuando el escenario se repite muchas veces (con este tamaño de muestra).

DOI: 10.20396/zet.v28i0.8657607

Este enfoque se relaciona con una analogía entre la medición de una probabilidad desconocida y la medición de cantidades físicas y revela una idea clave de probabilidad – la medición se llama estimación en jerga probabilística.

La medición de una cantidad desconocida es un proceso que usualmente se realiza mediante un procedimiento de medición bien definido. Sin embargo, en el contexto de probabilidad, se trata de cantidades no físicas que requieren métodos especiales que difieren de las mediciones físicas. La media de una población es una cantidad virtual que no existe en realidad, pero puede ser relevante conocerla. Lo mismo ocurre con una proporción de personas de una población que tienen una determinada propiedad; se trata de una cantidad virtual. O bien, el valor esperado desconocido de un modelo de probabilidad que se utiliza para describir la distribución de una variable (en una población finita o en un proceso de generación de datos). Asimismo, la probabilidad desconocida de un evento es “medida” por las frecuencias relativas del experimento aleatorio subyacente. Estas son cantidades virtuales. Sin embargo, la analogía con la medición de las magnitudes físicas puede aclarar el propósito y las propiedades de tales “mediciones”. De hecho, la medición de una cantidad física está relacionada con los errores de medición, el sesgo de la medición (un error sistemático de medición) y la precisión de las mediciones (los detalles de la analogía se encuentran en la Tabla 2).

En la teoría de las mediciones físicas, se ha recomendado promediar el resultado de unas pocas mediciones para disminuir el error de medición y aumentar la precisión del procedimiento de medición. Tales recomendaciones han sido establecidas por una conexión con la teoría de la probabilidad. Utilizamos la analogía de la medición de cantidades virtuales (como la media de una población o la proporción de personas en una población con una determinada propiedad) con la medición de cantidades físicas para mejorar la comprensión de los conceptos virtuales y su medición probabilística.

4.2 La tarea de medir – o estimar – una probabilidad desconocida

La idea es evaluar la precisión de varios instrumentos de medición comparándolos entre sí. En la Figura 3, vemos cómo la curva de las frecuencias relativas acumulativas se estabiliza con más ensayos; después de 1000 experimentos binarios las frecuencias relativas parecen ser muy estables. Para hacerlo más específico, investigaremos tres escenarios para investigar los 1000 datos. Podríamos usar el lanzamiento de monedas o cualquier otro experimento aleatorio binario; también podríamos imaginarnos centrarnos en uno de los nueve dígitos aleatorios del experimento de la Figura 1.

Tabla 2 – Comparación de los términos utilizados en la teoría de medición y en la probabilidad

<i>Categoría</i>	<i>Teoría de la medición</i>	<i>Teoría de la probabilidad</i>
Método	Medición de magnitudes físicas	Estimación de cantidades virtuales
Objeto de medida	El valor real de una entidad <i>física</i>	Media o proporción (u otros parámetros) de una población <i>real</i> o <i>virtual</i>

DOI: 10.20396/zet.v28i0.8657607

Proceso de medida	Para medir el objeto físicamente, use el instrumento repetidamente	De la población, simule datos de una muestra aleatoria
Resultado de medida	Algunos valores de medición	Datos de la muestra aleatoria
Valor de medida	El medio de los datos medidos	Una estadística conveniente de los datos de la muestra
Nombramiento del resultado	Un valor de medición del entidad física	Un valor de estimación de la característica de la población
Error de una medida	La diferencia entre el valor real de la entidad y el valor de medida	La diferencia entre el parámetro de la población y el valor estimado
Característica de una medida / error	El varía en general debido a condiciones incontrolables	El varía debido al carácter aleatorio en la generación de la muestra
Precisión de la medición	Talla general del error en general	Variabilidad de la estimación debido a aleatoriedad
Margen de error	Es medido en términos de la desviación estándar	Es medidos en términos de la desviación estándar
Calidad del procedimiento de medida	Más pequeño es el error estándar, mejor es el método de medición	Más pequeño el error estándar, mejor es el método de estimación

Fuente: Elaborado por el autor

- Escenario 1: Dividimos los 1000 datos (de 0s y 1s) en bloques de longitud 5 y usamos la frecuencia relativa en cada bloque (muestra) para “medir” (estimar) la probabilidad desconocida. Obtenemos 200 medidas (estimadas), que son significadas por cruces en la Figura 4 (izquierda); la representación muestra que – no inesperadamente – las medidas varían enormemente y bastantes medidas tienen un error de mayor igual a 0,20 (es decir, se encuentran fuera de la banda entre 0,30 y 0,70, la cual está marcada por gruesas líneas discontinuas).
- Escenario 2: Dividimos los 1000 datos en bloques de 10 de longitud y obtenemos 100 estimaciones de la probabilidad desconocida, que se representan en el diagrama central de la Figura 3. Mucho menos valores están más allá de la banda marcada para errores de medición mayor a 0,20.
- Escenario 3: Dividimos los 1000 datos en bloques de 20 de longitud y obtenemos 50 estimaciones de la probabilidad desconocida, que se muestran en el diagrama de la derecha de la Figura 3. Sólo dos mediciones tienen un error superior a 0,20.

Los cruces que representan las mediciones individuales de la cantidad desconocida muestran un patrón claro de reducción de izquierda a derecha en la Figura 3 con un número creciente de datos (longitud de la serie de mediciones, o tamaño de la muestra). También podemos comparar el efecto estabilizador de la duración del experimento en la curva de las frecuencias relativas (en la Figura 3) con la aleatoriedad completa a la que están expuestas las mediciones individuales de los bloques resultantes. Esto pone de relieve que no puede haber un efecto de compensación según el cual los siguientes resultados corrijan una desviación existente en las frecuencias relativas en la dirección del límite (la probabilidad). Este

pensamiento de compensación es muy popular, como puede verse en la ubicuidad de las listas de números recientemente sorteados en la lotería estatal, lo que proporcionaría la base para “calcular las probabilidades nuevas” de los números para el próximo sorteo de la lotería.

4.3 Análisis de la precisión de los tres instrumentos de medición

La analogía de la medición muestra que los instrumentos de medición difieren enormemente por la calidad de las mediciones; los errores del Instrumento 1 (Escenario 1) tienden a ser mucho mayores que los del Instrumento 3 (Escenario 3). Comparamos estos instrumentos de medición mediante un análisis estadístico de los errores de medición en nuestro escenario. Esto significa que comparamos el patrón de la distribución de la estimación de la probabilidad desconocida basada en datos repetidos de 5, 10 y 20 con un gráfico de barras (Figura 4). De nuevo, vemos el efecto de encogimiento. La muestra más grande ofrece una estimación más precisa de la probabilidad desconocida. Las Figuras 3 y 4 están ligeramente modificadas a partir de Borovcnik (2019a).

La Figura 4 prepara especialmente la idea de la distribución muestral de la proporción en una muestra de tamaño fijo. Es esencial aclarar explícitamente el carácter artificial de nuestro estudio de simulación de los instrumentos de medida: hemos repetido las mediciones para obtener una base de datos empírica sobre la calidad de los instrumentos. En la práctica, sin embargo, sólo tenemos *una* muestra y *una* medición (estimación) del parámetro desconocido. Aplicamos el conocimiento de tales estudios artificiales para transferir los conocimientos sobre la calidad a este único valor y así podemos juzgar su fiabilidad en el sentido de la dispersión (medida por la desviación estándar) de la distribución del muestreo de la proporción.

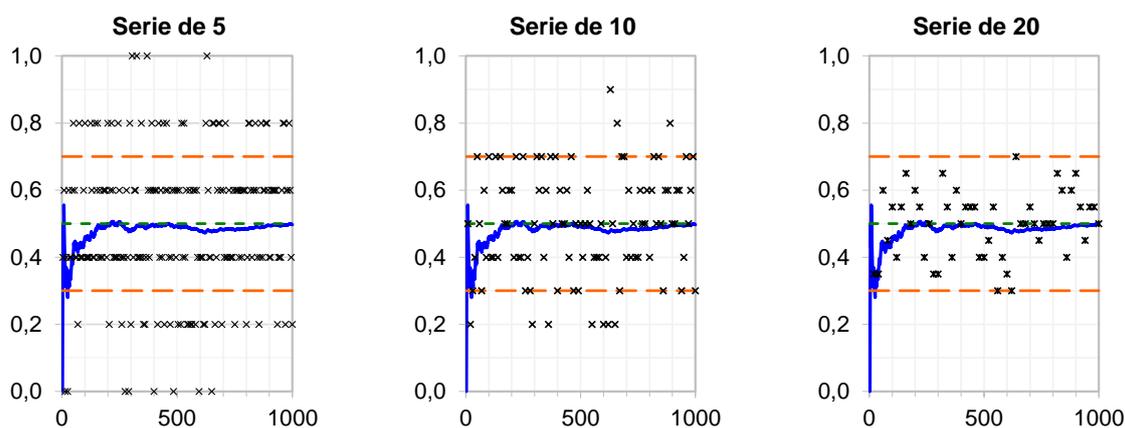


Figura 3 – “Convergencia” de las frecuencias relativas hacia la probabilidad desconocida por la duración de la serie de experimentos comparada con la medición repetida de esta probabilidad por diferentes procedimientos:

Izquierda: Medición de la probabilidad por la frecuencia relativa de los últimos 5 ensayos. En el medio:

Medición por el resultado de los últimos 10 ensayos. Derecha: Medido por los últimos 20 ensayos

La banda dentro de las líneas discontinuas marca las mediciones con un error menor de 0,20

Fuente: Elaborado por el autor

DOI: 10.20396/zet.v28i0.8657607

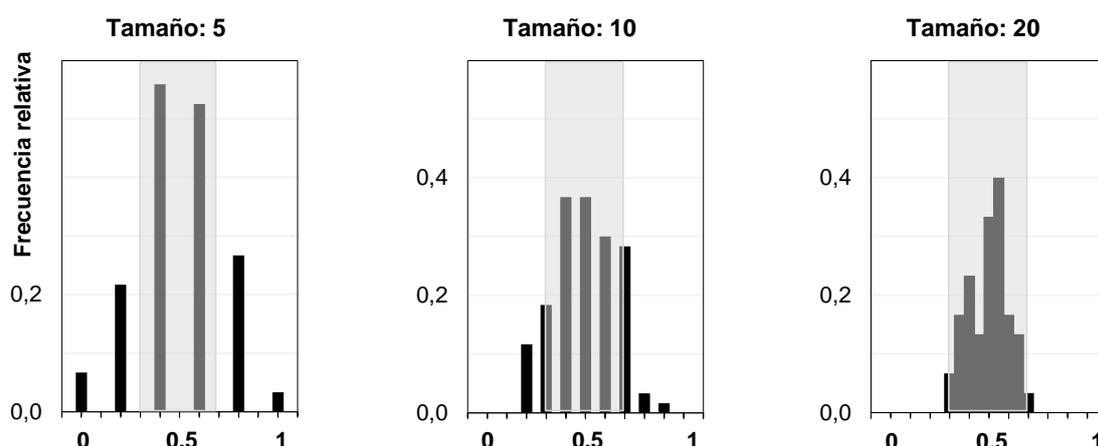


Figura 4 – Comparación de la precisión de los tres procedimientos para medir la probabilidad desconocida
Izquierda: Basado en los últimos 5 valores. En el medio: Basado en 10 valores. Derecha: Basado en 20 valores

El área sombreada significa un error de medición menor de 0,20

Fuente: Elaborado por el autor

Exploramos la precisión de los procedimientos de medición en lugar de investigar una ley de convergencia. Este experimento básico evita un comportamiento limitante oscuro de las frecuencias relativas y se centra en cambio en tres instantáneas que ilustran el fenómeno: un mayor tamaño de la muestra permite estimar la probabilidad desconocida con mucha más precisión. El experimento también muestra ya en la fase inicial para qué sirven las probabilidades y cómo podemos vincular la probabilidad y la estadística y cómo vamos a operacionalizar el concepto de información estadística y cómo podemos transferir las estimaciones de las muestras a la población (finito, o los procesos que generan las muestras).

Estudiar en el laboratorio estadístico garantiza que otros factores no podrían desdibujar los resultados, ya que tenemos control total sobre las suposiciones. La repetición del escenario de simulación (en las Figuras 3 y 4) es muy instructiva, ya que muestra un patrón estable en el efecto de reducción y permite a los alumnos experimentar su variación (cuándo cambia, cómo cambia, qué es invariable, cuál es el margen de variación o el error si se mide una cantidad) en lugar de investigar límites oscuros que no están abiertos al escrutinio.

El análisis de los 1000 datos en bloques de diferentes longitudes está motivado por un protocolo de lanzamiento de monedas. En una versión de simulación por ordenador se podría tratar de evitar que para las muestras de tamaño 20 haya muchos menos datos (sólo 50) en comparación con las muestras de tamaño 5 (donde tenemos 200 datos). Cuantos menos datos se tengan, mayor puede ser el efecto de confusión de la simulación. En una simulación dinámica se podrían simular 1000 (o más) muestras de tamaño 5, luego de tamaño 10, y finalmente de tamaño 20 para que las distribuciones empíricas sean menos propensas a artefactos de efectos aleatorios de la simulación y el patrón sea aún más claramente visible.

El experimento para promedios es análogo al actual. Sólo que la media en lugar de una proporción de una población debe ser medida por una muestra. Una vez más, se investiga

DOI: 10.20396/zet.v28i0.8657607

el procedimiento de medición. ¿Cuán precisa es la medición de la media de una población por la media de una muestra? En el laboratorio estadístico, podemos simular muestras de la distribución de la población para investigar nuestro dispositivo de medición. Es esencial recordar que en la práctica sólo se tiene una medición, ya que se tiene una sola muestra. El escenario de simular muestras repetidas de la población, determinar la media de esta muestra y continuar el muestreo con el cálculo de la media proporciona una base empírica para la distribución del muestreo de la media. Esta distribución del muestreo no es conceptualmente mucho más fácil en el escenario de la simulación, aunque se materializa en los datos sobre las mediciones repetidas que se generan en este proceso de muestreo; describe la calidad del instrumento de medida.

La distribución del muestreo se obtiene generalmente mediante teoremas matemáticos; constituye la clave para la inferencia estadística. Utilizamos el experimento introductorio a las frecuencias relativas para que los estudiantes se familiaricen con la conexión entre la probabilidad y la inferencia estadística. También es de interés investigar mediante un estudio de simulación qué tan rápida es la convergencia, es decir, cuántos datos de la muestra se necesitan para que la distribución se parezca bastante a una distribución normal. Una vez más, el teorema del límite central proporciona una descripción fácil de la distribución del muestreo de la media, ya que está aproximadamente distribuida normalmente.

4.4 Ampliación de la analogía de medición repetida al método Bootstrap

Es muy interesante que la analogía de la medición repetida dé una explicación natural para el método Bootstrap. Si conociéramos la población, podríamos simular una muestra tras otra, calculando siempre el parámetro de interés; este parámetro de la muestra proporciona una medida de la cantidad de población en la que estamos interesados. En Bootstrap, la simulación se realiza a partir de una aproximación de la población, la cual es proporcionada por la muestra original. Normalmente, se calcularía un intervalo de confianza sobre la base de esa muestra. Este intervalo de confianza da una información sobre dónde debe estar el parámetro “verdadero” de la población.

Aparte de los problemas conocidos en la interpretación de los intervalos de confianza, para parámetros distintos de la media de la población, el método proporciona fórmulas bastante complicadas para el intervalo de confianza. Una complejidad matemática que también puede ser la razón de los frecuentes malentendidos no sólo en las pruebas estadísticas sino también en los intervalos de confianza. Los intervalos de Bootstrap son mucho más fáciles de obtener, ya que sólo se necesitan los datos y sin más supuestos. El intervalo se alcanza simplemente mediante la simulación repetida de muestras de la primera muestra (con repetición). Sin embargo, la razón por la que el método funciona puede aclararse mucho mejor que usualmente por analogía con las mediciones.

Dados: los datos de una muestra de tamaño n con media y DE para una variable específica. ¿Cuán precisa es la media de la muestra como medida para el conjunto de la población? Si tenemos acceso a la población, podríamos tomar repetidamente una muestra de

DOI: 10.20396/zet.v28i0.8657607

ella y calcular el valor medio de esta muestra. Este valor medio representa la primera medición de la media “desconocida”. En el laboratorio estadístico (donde se conoce la población incluyendo su media), podemos investigar la precisión de nuestro procedimiento de medición para la media de la población mediante este método de medición de la media de la población por la media de la muestra.

Si generamos 1000 muestras (o incluso más) y calculamos la media, obtenemos una base empírica de medidas, que podemos analizar estadísticamente como en el caso de la tarea de medir una probabilidad desconocida (proporción de una población). Mediante este escenario de simulación, proporcionamos una base empírica para la distribución del muestreo de la media y podemos juzgar la calidad del procedimiento de medición (la estimación de la media “desconocida”). El procedimiento es exactamente el mismo que en el ejemplo de medición de una probabilidad desconocida.

Si no tenemos acceso a la población, no podemos muestrear más repetidamente de ella y medir (estimar) la media desconocida de la población. El siguiente procedimiento de medición, llamado Bootstrap, podría proporcionar una forma razonable de medir (estimar) la media desconocida de la población. En lugar de tomar una muestra de la población (lo que ya no es posible), tomamos una nueva muestra de los datos ya existentes (con repetición). Como este conjunto inicial de datos es una muestra aleatoria de la población, no debería haber demasiado error en este procedimiento de muestreo “aproximado”. Como en el caso de la población conocida, calculamos la media de esta muestra de Bootstrap y “medimos” (estimamos) la media desconocida de la población.

La calidad de este procedimiento de medición se evalúa repitiendo la medición, es decir, repitiendo para tomar una muestra de Bootstrap de los datos y calcular la media. Si el proceso se repite 1000 veces (o más), se obtiene una base empírica para la distribución de la media (ver Figura 6). Podemos reconocer cuánto varían estas mediciones de Bootstrap de una muestra a otra y encontrar un intervalo central (del 95 %, por ejemplo) de estas mediciones. Este llamado Bootstrap proporciona una evaluación del procedimiento de medición. Tiene propiedades similares a las de un intervalo de confianza para la media, pero tiene una interpretación diferente.

De hecho, es didácticamente atractivo lo simple que es la interpretación de un intervalo de Bootstrap del 95%: es ese intervalo central, que contiene el 95% de todas las mediciones de Bootstrap del parámetro desconocido de la población. Si el método de medición no tiene errores sistemáticos, entonces las mediciones repetidas deben variar alrededor del valor real. Y el intervalo Bootstrap indica la variabilidad de las mediciones y es el mejor que se puede obtener con este método, que tiene un margen de error inherente. La analogía con la medición de cantidades encaja perfectamente.

Desde principios de la década de los 90 hubo sugerencias para utilizar el Bootstrap y las estadísticas no paramétricas como una etapa transitoria en la enseñanza de la inferencia estadística con el fin de simplificar el primer acercamiento a este sofisticado tema. Este intervalo de Bootstrap es una imitación ingenua pero muy interesante del intervalo de

DOI: 10.20396/zet.v28i0.8657607

confianza clásico aunque Borovcnik (2019a) da razones por las que el método de los intervalos de Bootstrap no es adecuado para reemplazar los intervalos de confianza clásicos; podría utilizarse como una fase intermedia en la enseñanza hasta que se puedan enseñar los intervalos de confianza más sofisticados. Cobb (2007) ha sugerido reemplazar completamente la inferencia estadística por métodos de remuestreo, que incluyen Bootstrap y rerandomización. En la práctica, los expertos sabrían cuándo utilizar qué método para estimar la media de una población.

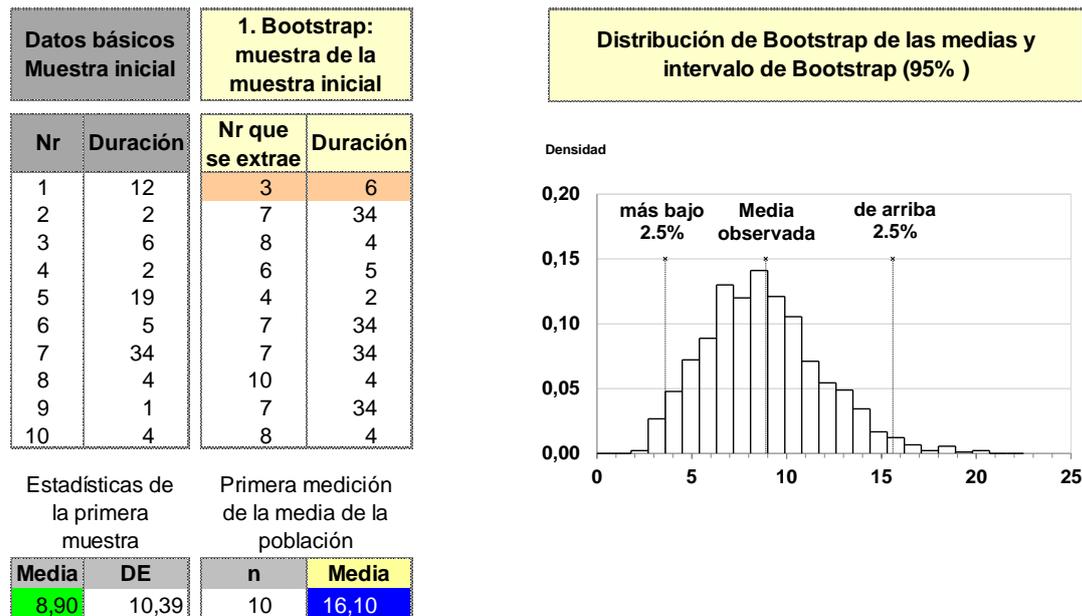


Figura 5 – Las muestras de Bootstrap para medir la media desconocida.

Izquierda: El primer Bootstrap muestra que la medición de la media desconocida es muy poco fiable.

Derecha: La distribución Bootstrap produce el 95 % de las mediciones repetidas entre 3,80 y 15,60

Fuente: Elaborado por el autor

5 Conclusiones y perspectivas

En el segundo capítulo, investigamos el carácter de probabilidad del experimento con dígitos aleatorios. La equiprobabilidad de los dígitos y el desarrollo de frecuencias relativas son dos partes inseparables del concepto y establecen una relación de complementariedad para la probabilidad.

Más bien que seguir consideraciones de convergencia obscuras, mostramos por nuestro experimento que el margen de fluctuación de las frecuencias relativas de los dígitos es mucho más pequeño para 1000 dígitos aleatorios que es para 50. Esto es un descubrimiento experimental. El resto, que el margen de fluctuación de 1000 dígitos a un tamaño mayor de los dígitos simulados continuará reduciéndose de manera similar, se debe a un experimento de pensamiento. Esto es típico para nuestras animaciones que sólo tenemos simulaciones para tamaños pequeños y establecemos un modelo en una comparación visual de los diagramas del experimento para varios escenarios. Entonces repetimos la simulación

DOI: 10.20396/zet.v28i0.8657607

para investigar si la forma ya encontrado permanecerá estable. La repetición de la simulación se puede seguir como ver un vídeo.

En el tercer capítulo, investigamos la complementariedad entre probabilidad y riesgo. Concentrarse en el aspecto de riesgo por tareas convenientes facilita entender el papel de la probabilidad. En el contrato de seguro hay un precio del cambio de las posiciones de riesgo (el cliente al principio) y certeza (la compañía de seguros al principio), que sólo se puede claramente determinar si uno quiere estimar, presuponer, o determinar subjetivamente el potencial del acontecimiento arriesgado para ocurrir. Qué número se atribuye a esta probabilidad puede no ser fácil a solucionar, pero un escenario con relación a este número aclara qué decisión se puede preferir: sacar el seguro o no sacar.

De manera interesante, el ejemplo también toca la complementariedad entre frecuencias y los aspectos subjetivistas de la probabilidad ya que la compañía de seguros puede usar las frecuencias del acontecimiento asegurado (p.ej., el accidente de un coche) mientras el cliente sólo tiene consideraciones cualitativas (subjetivistas) para encontrar una probabilidad conveniente para este acontecimiento. En la tarea de la revista, hay soluciones matemáticas sin probabilidades (busca al mínimo la pérdida máxima, p.ej.); aún, tales estrategias pueden llevar a soluciones inaceptables como en el ejemplo con las copias de la revista. El papel virtual de probabilidades en relación con riesgos se hace mucho más claro por tales tareas.

En el cuarto capítulo, introdujimos consideraciones de la inferencia informales en las fases tempranas de enseñar la probabilidad. Estas consideraciones hacen el concepto de la probabilidad mucho más accesible y natural. En lugar de “mostrar” un límite de frecuencias relativas hacia una probabilidad desconocida, uno analizaría la tarea de estimar la probabilidad desconocida por la frecuencia relativa de una serie de experimentos con una talla fija. Entonces desarrollamos la analogía con la medida de entidades físicas e investigamos la precisión de nuestro instrumento de medición.

El análisis de los instrumentos de medida proporciona el conocimiento: el instrumento con el conjunto de datos más grande es más preciso. El resto se hace con un experimento del pensamiento: cuantos más datos, más preciso es el instrumento de medición. La analogía con la medida de una cantidad física también da una explicación excelente por el método de Bootstrap, que ha sido adoptado por la escuela de “Inferencia Informal” en la educación de estadística

La probabilidad es un concepto virtual a pesar de todos los esfuerzos de ilustrar su impacto por la simulación de la situación aleatoria investigada. El concepto tiene muchos aspectos que son significados por una especie de complementariedad: Equiprobabilidad y frecuencias relativas, frecuencias y aspectos subjetivistas de probabilidad, probabilidad y riesgo, y probabilidad e inferencia estadística. La complementariedad significa que la separación de los aspectos dejaría el concepto con la pérdida completa de su sentido genuino. Como una consecuencia, simplificando las cuestiones con el objetivo didáctico puede deformar el sentido y trae la educación de la probabilidad en un dilema.

Como la probabilidad es un concepto teórico, la gente carece de experiencia en ella, lo que puede corregir intuiciones básicas inadecuadas. La interacción entre estas intuiciones primarias e intuiciones secundarias se vuelve relevante para establecer concepciones estables en los alumnos. Borovcnik (2019b) proporciona un marco teórico para identificar ideas fundamentales de probabilidad y cómo desarrollar intuiciones que sean estables. Este marco se refiere a la interacción de Fischbein entre las intuiciones primarias y secundarias.

Desarrollar intuiciones secundarias mediante tareas matemáticas resulta más fácil si el diseño de las tareas y los itinerarios de aprendizaje previstos se centran en el propósito de la probabilidad. Esto ayuda no sólo a motivar a los estudiantes para que continúen su trabajo, sino que también puede convencerlos de que los conceptos son útiles, pero lo más importante es que el propósito, una vez sacado a la luz, permite que las matemáticas se desarrollen con un objetivo para que el concepto empieza a tener sentido. Al principio, tiene sentido ya que ayuda a resolver la tarea. Entonces, tiene sentido, ya que en el contexto de la tarea y en relación con el propósito de la tarea, el concepto puede volverse natural con propiedades obvias, o al menos con propiedades que parecen razonables (y no contrarias a la intuición). Trabajando repetidamente en tales tareas con un propósito explícito, los estudiantes pueden completar su red cognitiva para ese concepto y obtener una comprensión más amplia. Un entendimiento como el de los ingenieros a veces se desarrolla: sin conocer todos los detalles matemáticos, tienen un buen entendimiento que les permite sacar sus conclusiones.

Referências

- Batanero, C. & Borovcnik, M. (2016). *Statistics and probability in high school*. Rotterdam: Sense.
- Batanero, C., Chernoff, E., Engel, J. Lee, H., & Sánchez, E. (2016). *Research on teaching and learning probability. ICME-13 Topical Surveys*. Cham, Switzerland: Springer online. Online: doi.org/10.1007/978-3-319-31625-3_1.
- Batanero, C., Henry, M., & Parzysz, B. (2005). The nature of chance and probability. In: A. G. Jones (Ed.), *Exploring probability in school* (pp. 15-37). New York: Springer.
- Bennett, D. J. (1999). *Randomness*. Cambridge, MA: Harvard University Press.
- Bohr, N. (1934/1927): *Atomic theory and the description of nature*. Cambridge, UK: Cambridge University Press (containing the *Como lecture* of 1927).
- Borovcnik, M. (1992). *Stochastik im Wechselspiel von Intuitionen und Mathematik (Stochastics in the interplay between intuitions and mathematics)*. Mannheim: Bibliographisches Institut.
- Borovcnik, M. (2015a). Central theorems of probability theory and their impact on probabilistic intuitions. In J. M. Contreras, et al. (Eds.), *Didáctica de la Estadística, Probabilidad y Combinatoria, 2* (pp. 15-35). Granada, Universidad de Granada.
- Borovcnik, M. (2015b). Risk and decision making: The “logic” of probability. *The Mathematics Enthusiast*, 12(1, 2 & 3), 113-139.

DOI: 10.20396/zet.v28i0.8657607

- Borovcnik, M. (2016). Probabilistic thinking and probability literacy in the context of risk. *Educação Matemática Pesquisa*, 18(3), 1491-1516.
- Borovcnik, M. (2019a). Inferência informal e inferência “informal”. *Educação Matemática Pesquisa*, 21(1), 433-460. Online: [dx.doi.org/10.23925/1983-3156.2019v21i1p433-460](https://doi.org/10.23925/1983-3156.2019v21i1p433-460).
- Borovcnik, M. (2019b). Conceptos fundamentales y sus propiedades clave en probabilidad – Cómo identificarlos y proporcionar intuiciones que los sostengan. *Revista Eletrônica de Educação Matemática (REVEMAT)*, 14, *Educação Estatística*, 1-25.
- Borovcnik, M., & Bentz, H. J. (1990, 2003). Intuitive Vorstellungen von Wahrscheinlichkeitskonzepten: Fragebögen und Tiefeninterviews (Intuitive conceptions of probabilistic concepts: Questionnaire and in-depth interviews). Technical Reports. Klagenfurt University.
- Borovcnik, M., & Bentz, H. J. (1991). Empirical research in understanding probability. In R. Kapadia, & M. Borovcnik (Eds.), *Chance encounters. Mathematics education library vol 12* (pp. 73-106). Dordrecht: Kluwer Academic Publishers.
- Borovcnik, M., & Kapadia, R. (2014). A historical and philosophical perspective on probability. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic thinking: Presenting plural perspectives* (pp. 7-34). New York: Springer.
- Borovcnik, M., Peard, R. (1996). *Probability*. In A. Bishop, et al. (Eds.), *International handbook of mathematics education* (pp. 239-288). Dordrecht: Kluwer.
- Carranza, P. & Kuzniak, A. (2008). Duality of probability and statistics teaching in French education. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education*. Monterrey: ICMI and IASE.
- Chernoff, E. J. (2013). Probabilistic relativism: A multivalentological investigation of normatively incorrect relative likelihood comparisons. *Philosophy of Mathematics Education Journal*, 27, 1-30.
- Chernoff, E. & Sriraman, B. (Eds.) (2014). *Probabilistic thinking: presenting plural perspectives*. Advances in Mathematics Education (Vol. 7). New York: Springer.
- Cobb, G.W. (2007). The introductory statistics course: A Ptolemaic curriculum? *Technology Innovations in Statistics Education*, 1(1).
- de Finetti, B. (1937/1992). La prévision: ses lois logiques, ses sources subjectives. *Annales Institut Henri Poincaré*, 7, 1-68 (1937); English translation: Foresight: Its logical laws, its subjective sources. In S. Kotz, & N. L. Johnson (Eds.), *Breakthroughs in statistics* (pp. 134-174). New York, Berlin: Springer.
- Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. Dordrecht: Reidel.
- Fischbein, E. (1987). *Intuition in science and mathematics*. Dordrecht: Reidel,
- Freudenthal, H. (1972). “The empirical law of large numbers” or “The stability of frequencies”. *Educational Studies in Mathematics*, 4(4), 484-490.
- Gigerenzer, G. (2002). *Calculated risks: How to know when numbers deceive you*. New York: Simon & Schuster.

DOI: 10.20396/zet.v28i0.8657607

- Jones, G. A., Langrall, C.W., & Mooney, E. S. (2007). *Research in probability: Responding to classroom realities*. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*, Vol. 2 (pp. 909-956). Greenwich, CT: National Council of Teachers of Mathematics, Information Age Publishing.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kapadia, R., Borovcnik, M. (Eds.) (1991). *Chance encounters. Probability in education*. Dordrecht: Kluwer.
- Kolmogorov, A. N. (1933/1956). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin; English translation: *Foundations of the theory of probability*. New York: Chelsea.
- Lecoutre, M. P. (1992). Cognitive models and problem spaces in “purely random” situations. *Educational Studies in Mathematics*, 23(6), 557-568.
- Simons, P. R. (1984). Instructing with analogies. *Journal of Educational Psychology*, 76(3), 513–527.
- Spiegelhalter, D. (2014a). Comments on probabilistic thinking. In: E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic thinking: presenting plural perspectives* (Back cover). New York: Springer.
- Spiegelhalter, D. (2014b, April). *What can education learn from real-world communication of risk and uncertainty?* Invited lecture at the Eight British Congress on Mathematical Education, Nottingham.
- Steinbring, H. (1991). The theoretical nature of probability in the classroom. In: R. Kapadia & M. Borovcnik (Eds.). *Chance encounters* (pp. 135-167). Dordrecht: Kluwer.
- Székely, G. J. (1986). *Paradoxes in probability and mathematical statistics*. Dordrecht: D. Reidel.
- Vancsó, Ö. (2009). A parallel discussion of classical and Bayesian ways as an introduction to statistical inference. *International Electronic Journal of Mathematics Education*, 4(3), 291–322.
- Vancsó, Ö. (2018, July). How visualisation using software helps understanding classical and Bayesian statistics. Invited paper in session 6C “Teaching Probability in School – Understanding and Linking it to Statistics.” ICOTS 10, Kyoto, 8-13 July, 2018. Online: www.researchgate.net/profile/Oedoen_Vancso.
- von Plato, J. (1994). *Creating modern probability: its mathematics, physics and philosophy in historical perspective*. Cambridge: Cambridge University Press.