UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO
DOCTORADO EN CIENCIAS (FÍSICA)

DE LA PROBABILIDAD, LA MECÁNICA CUÁNTICA, EL TEOREMA DE BELL Y OTRAS VANIDADES DE LA CIENCIA

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS (FÍSICA)

PRESENTA:
ALDO FERNANDO GUADALUPE SOLIS LABASTIDA

DIRECTOR DE TESIS:
DR. JORGE GUSTAVO HIRSCH GANIEVICH
INSTITUTO DE CIENCIAS NUCLEARES, UNAM

COMITÉ TUTOR
DR. ALFRED BARRY U'REN CORTÉS
INSTITUTO DE CIENCIAS NUCLEARES, UNAM

DR. ELÍAS OKON GURVICH
INSTITUTO DE INVESTIGACIONES FILOSÓFICAS, UNAM

CIUDAD UNIVERSITARIA,CD. MX., ENERO 2024

De la Probabilidad, la Mecánica Cuántica, el Teorema de Bell y otras Vanidades de la Ciencia

Aldo Fernando Guadalupe Solis Labastida

2023

Índice general

Parte I

No existe “La teoría de la probabilidad”

Capítulo 1

Yo confieso

De los textos de fundamentos de la mecánica cuántica que he leído no hay uno que encuentre más excitante que el prólogo del libro de Pitowsky sobre probabilidad[2]. El libro de hecho no habla sobre probabilidad cuántica, lo que sea que eso signifique, sino más bien sobre una serie de reflexiones y exploraciones de él en la descripción de fenómenos cuánticos usando probabilidad.

En mis propias palabras el artículo transmite una chispa de sentido común, nos enfrenta a una cuestión muy natural que siempre ha estado ahí quieta, invisible, que parece una locura no haberla visto antes. El huevo de Colón de los fundamentos de la mecánica cuántica.

Si bien hoy día, y en días jóvenes de Pitowsky, la desigualdad de Bell, el teorema de Kochen-Specker, etc. ya parecían dar la victoria a la mecánica cuántica sobre una descripción más clásica; esa victoria parece muy fácilmente obtenida considerando eventos similares en historia de la ciencia.

Es decir, de alguna manera, se convirtió en opinión general, que estos teoremas prohíben la creación de una teoría "realista y local", lo que sea que eso signifique. Una teoría que asigne valores a las variables de un experimento tipo Bell y que al mismo tiempo cumpla ciertas nociones de localidad. Eso no me parecía para nada descabellado en el momento de aprender sobre las desigualdades. De hecho me pareció muy natural y, haciendo gala de mi pensamiento crítico, pensé que la conclusión sobre la no localidad era ineludible. Me gustaría sentir culpa por dicha actitud, pero después de leer a ganadores del Nobel como Clauser en misma línea, me parece natural que mi educación me lleve a tomar los dogmas de la ciencia.

Cuál es entonces la cuestión que nos presenta Pitowsky. Si es cierto que los teoremas no-go demuestran la imposibilidad de realizar dicha teoría; con qué facilidad han obtenido dicha victoria ¿no lo cree usted? De hecho parecería una proeza digna de poema épico realizar un trabajo como el del teorema Kochen-Specker o los primeros experimentos de Bell. Pero concordará usted que el teorema de Bell consiste en algunas pocas líneas en las que se calculan no más que valores esperados con probabilidades condicionales. Concordará usted también que el teorema de Kochen-Specker, también se le puede meter en un par de cuartillas, aunque claramente encontrar el conjunto de vectores con las propiedades adecuadas no es problema trivial. Es decir, dichos teoremas son matemáticamente muy simples lo que se contrapone a la grandiosidad del resultado que afirman demostrar. Pitowsky afirma, según mi exégesis, que para el ojo de quien se ha entrenado en la matemática puramente formal, de hecho dichas aseveraciones parecen sospechosas.

Dentro de la matemática moderna hay ya toda una serie de teoremas y principalmente contraejemplos que desafían hasta al más esforzado de los adictos al sueño hilbertiano. El teorema de la bola de Banach, las dimensiones fraccionarias de Peano y demás objetos han enseñado al matemático moderno y esforzado a ser pulcro en sus hipótesis, sus conceptos y sobre todo a sujetar muy bien su lengua al hacer afirmaciones. De alguna manera estos fenómenos hacen saltar una duda sobre los teoremas no-go.

A qué matemático moderno no le parece sospechoso que unas cuantas líneas de probabilidad, o un francamente básico estudio de las propiedades de las bases en R^3 tenga como consecuencia que no se pueda hacer una teoría realista local. Cómo, si hay objetos tan raros como la bola de Banach, resulta que esto es suficiente para limitar un objeto tan complejo como una teoría.

Con esto viene la comparación obligatoria con uno de los teoremas más importantes en la historia de las matemáticas, al menos su parte moderna: el teorema de Gödel. Éste es una de las contadas aseveraciones matemáticas tan potentes, o ingenuas, como para atreverse a prohibir teorías.

Esos teoremas requieren un contexto bastante complicado para plantearse. Una gran cantidad de definiciones: símbolo, lenguaje formal, sintaxis, etc. Sumado a esto tenemos una ejecución matemática que eriza la piel: expresar la aritmética con los números mismos es una idea que justifica que Gödel sea comparado en la lógica con nombres como Aquino y Aristóteles. Agreguemos a esto una "manufactura matemática" de un filo que parece podría partir a un diamante en dos. Qué habilidad necesaria para ingeniar la función beta, combinar con el método diagonal y llegar a la famosa sentencia que permite decir cosas sobre su propio número de Gödel. Y todo para poder constreñir lo que una teoría aritmética recursiva puede o no puede hacer[38].

Si para limitar a las teorías aritméticas se ha necesitado semejante obra del pensamiento humano, parece extraño pensar que unas cuántas líneas de probabilidad o un estudio relativamente básico de las propiedades de las bases en R^3 puedan realmente prohibir todas, e insisto, todas las teorías realistas locales ¿no le parece? Cómo es posible que con tan poco sean capaces de realizar una tarea que toma básicamente un libro entero en el caso de Gödel[1]

Esa es la cuestión de Pitowsky, es más bien una duda, una duda razonable, una duda pertinente. En vista de la obra matemática actual; en vista de los contraejemplos en topología a los que se les dedica libros[73]; en vista de las mil veces que los mejores matemáticos han visto engañado su intuición en las formas más impresionantes; en vista que probar el teorema de Gödel requiere de una obra magna de habilidades matemáticas sobrehumanas ¿realmente podemos creer que unas cuantas líneas que por comparación no tienen el mínimo de complejidad son capaces de limitar que existan teorías realistas locales que describan los fenómenos cuánticos?

Al intentar coincidir con Pitowsky uno se encuentra con que la simplicidad de los teoremas no-go es su fortaleza. Al ver la ecuación que lleva a los teoremas de Bell uno no puede encontrar mucho donde objetar. Como ya dijimos son solo algunos valores esperados de probabilidades condicionales, cómo no convencerse con un argumento tan sólido y simple.

Fue solo hasta ver que todas estas ecuaciones tienen una cantidad horrible de P's al escribirse apropiadamente, que me percaté de que no entiendo qué significa esa letra, esa función, esa medida. Me percaté que todos los días hablo de probabilidades de llover o de ganar de la misma forma que hablo de probabilidades de que "un electrón dé arriba al medir su espín". ¿En serio es tan relevante y poderosa la teoría de la probabilidad para abarcar semejante campo de aplicación? Pocas son las teorías matemáticas que le pueden competir y tendrían que ser ya en el campo de las más grandes digamos la aritmética o el cálculo.

Tal como el aire que nos rodea y sin el cual no podríamos vivir, la probabilidad se ha visto indispensable para el discurso científico de hoy. Es esta herramienta la que ha permitido que la ciencia llegue a lugares con los que antes no tenía nada en común y pocas cosas quedan en la ciencia que no se apoyen en ella de un modo u otro.

Entre Dios, usted y yo, debo confesar que sin importar mi preparación

[1]En general los libros de lógica matemática a un nivel universitario tienen dicho teorema al final de sus páginas donde ya han desarrollado las herramientas necesarias para atacar el tema[44].

como científico, llevo años hablando de probabilidades con una confianza loca en que lo que digo es cierto, pero no sé de probabilidad. Cursé varias clases acerca de la probabilidad y me dediqué esforzadamente a hacer los ejercicios del libro, a tener buenas calificaciones y tener la respuesta que los maestros buscaban, pero nunca entendí el tema. Mil veces me llené la boca con discusiones sobre los axiomas de Kolmogorov y por qué se deben aplicar en una situación de cierta manera, pero nunca me tomé la molestia de leer a Kolmogorov.

Cómo he pasado años convencido de que los axiomas de Kolmogorov capturan eso que llamamos probabilidad sin siquiera haberlos leído de él, no lo sé. Pero lo mismo podría decir de mi conocimiento sobre la física. Me he dicho que sé mecánica, me convencí de eso luego de aprobar varios cursos donde se discuten de nivel básico a avanzado sus diferentes formulaciones. ¿Entiendo acaso el primer escolio de los Principia? No. En algún momento me pregunté qué era un sistema de referencia inercial, tampoco. Ya no digamos si me pregunté cómo se pueden seguir utilizando las tres leyes de Newton una vez que se quitó el espacio absoluto. Es mucho pedirle a un estudiante como yo cuya llama del pensamiento crítico era débil y se fue apagando con la costumbre de respetar dogmáticamente los libros de texto y hacer como penitencia los problemas del libro.

Así que la discusión sobre la mecánica cuántica, los teoremas de Bell y la probabilidad deben pasar forzosamente por esta confesión donde le comparto que he pecado mucho de vanidad. Vanidad de pensar que simplemente por tener una preparación en la física yo tengo alguna idea de lo que es la probabilidad. Pues sin importar la ingente cantidad de cálculos que he hecho sobre valores promedio, distribuciones, verosimilitudes y demás; solo he hecho cálculos, solo he escrito símbolos en un papel o en una computadora y, sin importar si podía reproducir el teorema del límite central de memoria, para eso no se necesita pensar, no se necesita reflexionar. Solo se necesita escribir eso que todos llamamos lo "correcto".

Este es un pecado que llevo en la probabilidad, la mecánica cuántica, los teoremas de Bell y un sin fin de temas que me avergüenza listar. Le pido me perdone la vanidad de escribir sobre estos temas.

Capítulo 2

No existe solo una teoría de la probabilidad

Uno de los objetivos principales de este trabajo es poner de relieve la importancia que tienen los fundamentos de la probabilidad para la mecánica cuántica y su interpretación. Por lo cual, abordaremos en primer lugar a la probabilidad como un objeto más complicado de lo que su práctica diaria podría demostrar. Las personas tenemos una visión de la probabilidad que viene de la práctica y tiene componentes de escuelas de pensamiento anteriores con las que no tenemos mucha familiaridad. Los discursos de éstas no siempre se articulan bien y a veces chocan entre sí. Sin embargo, es difícil apreciar esto pues la probabilidad se ha vuelto muy cotidiana. Es una de esas palabras que el uso las ha gastado tanto que ya parecen no significar nada.

2.1. Un lugar para la probabilidad

Hoy, la probabilidad parece indispensable para la comunidad científica. La validez de un estudio viene justificada por el uso de técnicas basadas en estadística y probabilidad. Una buena parte de los nuevos artículos o reportes científicos que lidian con datos experimentales, al producirlos o analizarlos usan método estadísticos. En general la repetibilidad de un resultado es uno de los criterios más deseables en la comunidad científica. Cualquiera de estos trabajos está lleno de valores promedio, desviaciones estándar, valores P y otras cantidades relacionadas. En particular, el valor P se ha convertido en la prueba de hipótesis en la prueba de hipótesis más usada, incluso al grado de considerarse un requerimiento para la publicación de artículos en algunas comunidades científicas[49]. Dentro de éstas, un estudio que no haya aplicado dichas técnicas tiene una desventaja sobre los que sí lo hacen, pues se puede llegar a considerar que no se ha "probado" por la evidencia y en el caso extremo se puede considerar que no es científico. En definitiva, estas técnicas ya se consideran por sí mismas un criterio de validación para el nuevo conocimiento científico.

Es natural pensar que esta situación viene acompañada de una visión crítica de la probabilidad y sus métodos. Después de todo, si estas técnicas tienen tanto peso en la ciencia debe de haber buenas razones para ello. Sin embargo, los fundamentos de la probabilidad, y por extensión de la estadística, son poco conocidos e ignorados por buena parte de la comunidad. Cuando se nos cuestiona a los científicos "¿qué es la probabilidad?", esto proviene de las personas que no pertenecen al gremio, mas no como parte de nuestra labor científica cotidiana.

Tal interrogante, cabe aclarar, no interfiere con esta labor; pues es posible aplicar las técnicas mencionadas sin cuestionar su fundamento. Una persona, científica o no, puede calcular promedios diariamente sin preguntarse qué significan, para qué se usan o si son pertinentes en esa situación. Por ejemplo, los valores P, que permiten decidir entre dos hipótesis, actualmente son calculados usando programas informáticos sin intervención humana. Es decir, se introducen los datos a la computadora y solo se espera el resultado positivo o negativo. Incluso la significancia de la prueba tiene un valor por defecto de 5 % que es desconocido para muchos usuarios[1].

Esta situación que no parece dañina en principio puede tener consecuencias graves, ya que algunas decisiones de Estado, con impacto en la vida de millones de personas, se han tomado con base en estos métodos. Por tanto, si queremos ser coherentes con la imagen de la ciencia que promueve la crítica, la fundamentación de la probabilidad tiene un papel central y la pregunta sobre la pertinencia de sus métodos es cotidiana. Éste es el contexto en que tiene lugar cuestionarse "¿qué es probabilidad?".

2.2. La visión cotidiana de la probabilidad en la ciencia

Los dados y las monedas, muy cercanos a los juegos de azar, parecen ser la primera imagen de la probabilidad. En realidad, fueron los juegos de azar los que dirigieron buena parte de los comienzos de la teoría de la probabilidad. La práctica cotidiana es relativamente simple: se cuenta el número de casos posibles, luego todos se consideran igual de probables y, debido a que la probabilidad debe sumar uno, la probabilidad de cada uno es 1 divido entre el número de casos. Por ejemplo, lanzar una moneda tiene dos posibles resultados y su probabilidad es 1/2. Para el dado tenemos 6 resultados con probabilidad 1/6 para cada uno. Este procedimiento en apariencia sencillo es la base del cálculo de probabilidades en los juegos, y en buena medida de la práctica cotidiana.

Las reglas antes mencionadas pueden dar lugar a resultados que no van de acuerdo con nuestra idea "cotidiana" de probabilidad. Vamos a aplicarlas de dos maneras diferentes, para demostrar esta situación:

Si lanzamos dos dados, ¿cuál es la probabilidad de que la suma de sus resultados sea 12?

1. Como cada dado tiene seis caras, hay 36 posibles resultados; por lo que la probabilidad para cada resultado es 1/36. Ahora bien, la suma sólo será 12 cuando ambos dados resulten 6, y entonces la probabilidad de este evento es de 1/36.

2. Otro argumento es que los resultados del juego van desde el número 2 hasta el número 12, es decir tenemos 11 posibilidades. Inmediatamente, asignamos a cada uno la misma probabilidad, esto es 1/11, concluyendo que el valor correcto debe ser 1/11.

Como podemos ver ambas argumentaciones concuerdan con las reglas de asignación de probabilidades, aunque esperamos un único valor para la probabilidad. Por otro lado, esperamos que coincida con los "datos", esto es, si tiramos dos dados muchas veces, la fracción de veces que sale doce, llamada la frecuencia relativa, debe tener un valor parecido al de la probabilidad. De esta manera, tenemos una forma "correcta" de asignar probabilidades, pues la frecuencia relativa será muy cercana a 1/36.

Esta condición le da a la probabilidad una imagen más objetiva, pues depende de los resultados de algo que no está en nuestra voluntad. En ese caso, a veces se dice que tenemos una herramienta con gran poder predictivo. Por un lado, contamos los posibles casos, lo que es una práctica completamente simbólica. Por otro, determinamos de forma empírica el valor de la frecuencia relativa que es una cantidad "experimental". Así la probabilidad parece ser una ciencia empírica con su parte "teórica" y su parte "experimental". Sin embargo, esta situación se puede llevar al extremo: si ambos números deben coincidir, no vale la pena hablar del conteo de casos por un lado y las frecuencias relativas por otro, son lo mismo. No es difícil encontrar a personas dedicadas a la ciencia aseverando que la probabilidad es por definición la frecuencia relativa.

Al mismo tiempo, si ya establecimos que la probabilidad tiene que coincidir, dentro de cierto intervalo, con estas frecuencias relativas, bien podemos dedicarnos a medirlas y asumir que este es el valor de la probabilidad y hasta asociarle una incertidumbre como a las cantidades físicas. Si por alguna razón no tenemos la posibilidad de contar casos, por desconocimiento o dificultad, esta parece una excelente alternativa para determinar las probabilidades. De hecho, así se construyen una gran parte de las probabilidades hoy día en todas las ciencias.

Tomemos por ejemplo la probabilidad de lluvia que es bastante cotidiana no solo para los científicos sino para un público más amplio. En primer lugar, contar los casos no da un único resultado, pues hay varias formas de conteo como usar con el número de gotas que van a caer, los milímetros de lluvia, la temperatura o el número de nubes. Luego, la asignación de probabilidades pasa a las frecuencias relativas. Para calcularlas se puede usar un histórico de la lluvia que esté clasificado según las distintas condiciones en las que se dio: presión, temperatura, etc. Se cuentan el número de veces que llovió en condiciones "similares" y se divide entre el total. Así se puede obtener una "probabilidad de lluvia".

Desgraciadamente, la solución no es completa pues tener un histórico o una base de datos puede llevar preguntas como: ¿debo contemplar el día de la semana al asignar probabilidades? ¿el color de los zapatos que usé? Y aunque se pueden articular razones para tomar o descartar variables, también es de subrayar que no hay un cálculo independiente de los "datos", que de alguna manera desdibuja la imagen de "teoría y experimento" de la que hablábamos antes.

Aún así, el principal problema del histórico es que no existe para una gran cantidad de situaciones, no solo es desconocido sino que de principio es imposible tenerlo. Pongamos por caso la pregunta ¿cuál es la probabilidad de que nuestro equipo gane este partido de pelota, dado que es la primera vez que jugamos? Es claro que no puedo realizar muchos juegos, contar el número de veces que ganamos y dividirlo entre el total simplemente porque es la primera vez que jugamos. También es bastante claro que contar casos nos llevaría al mismo problema que con la lluvia. En conclusión, no tenemos forma de obtener el valor de probabilidad. Además situaciones análogas a esta son parte de la vida diaria, solo hace falta mirar a la infinidad de apuestas que se hacen alrededor del mundo de los deportes.

Todos los elementos mencionados se encuentran en la visión cotidiana de la probabilidad con la que, en una medida u otra, las personas somos familiares. Es una visión que quiere mirar a la probabilidad como una cantidad que es objetiva pero al mismo tiempo su cálculo requiere elementos subjetivos, quiere aplicarse a casos únicos aún cuando no haya datos que la sustenten, quiere aplicarse a toda situación sin ataduras o límites.

Manejar nociones incoherentes de un concepto no tiene en sí nada de raro. Todos lo hacemos de manera cotidiana. Sin embargo, la ciencia tiene una imagen pública que sostiene la coherencia como una de sus principales características. Así que, de querer ser coherente, será necesario un concepto de probabilidad más robusto. También es posible ignorar la coherencia o simplemente mirar a otro lado que es de hecho, históricamente, la opción preferida.

2.3. Sobre lo que hoy llamamos los axiomas de Kolmogorov

Empecemos por establecer a qué llamamos hoy axiomas de Kolmogorov. Hoy se asume en buena parte de la comunidad científica que la probabilidad está completamente caracterizada por dichos axiomas. Se cree, de manera ingenua, que esta descripción es acabada, que cualquier problema que concierna a la probabilidad está determinado. Esto ignora la historia de la probabilidad y a las muchas escuelas de pensamiento que sobre ella han reflexionado.

Los axiomas de Kolmogorov, no son estrictamente lo que Kolmogorov escribió originalmente. En la forma en la que hoy día se pueden encontrar en los libros de texto [17, 39, 68], por poner un ejemplo representativo de estos axiomas veamos la manera en la que son expresados en Wikipedia en inglés[5]:

> Let (Ω, F, P) be a measure space with $P(E)$ being the probability of some event E, and $P(\Omega) = 1$. Then (Ω, F, P) is a probability space, with sample space Ω, event space F and probability measure P.
>
> The probability of an event is a non-negative real number:
>
> $$P(E) \in \mathbb{R}, P(E) \geq 0 \qquad \forall E \in F$$
>
> where F is the event space. It follows that $P(E)$ is always finite, in contrast with more general measure theory. Theories which assign negative probability relax the first axiom.
>
> Second axiom This is the assumption of unit measure: that the probability that at least one of the elementary events in the entire

sample space will occur is 1

$$P(\Omega) = 1.$$

Third axiom. This is the assumption of σ-additivity:

Any countable sequence of disjoint sets (synonymous with mutually exclusive events) $E_1, E_2, \ldots$ satisfies

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

Hablemos con un poco más de detalle de la terna (Ω, F, P). A Ω solo se le pide ser un conjunto matemático y se le llama, en español, el *espacio muestral*. En segundo lugar tenemos a F, un conjunto, cuyos elementos son subconjuntos de Ω, con ciertas propiedades de cerradura y cuyo nombre común es el *espacio de eventos*. Finalmente, tenemos la función de P que tiene como dominio el espacio de eventos y toma valores en los números reales no negativos. Esta función P es a la que se le llama *medida de probabilidad* o simplemente *probabilidad*. Esta nomenclatura es la que se puede encontrar en la mayoría de los textos de probabilidad actuales.

La comunidad dedicada a la probabilidad, dice sin especificar los detalles matemáticos, que las distintas escuelas de la probabilidad convergen en los axiomas de Kolmogorov[36]. Es cierto, para esos ejemplos básicos de libro de texto. Cada escuela, sin importar lo distinto de su aproximación, cumple dichos axiomas, al menos en los casos más comunes. Con esto en mente pareciera que la relación entre la aplicación de la probabilidad y su fundamentación tuvieran poco que discutir; como si los axiomas de Kolmogorov fueran una especie de interfaz perfecta que divide los problemas de la aplicación y los de la fundamentación permitiendo una perfecta división del trabajo probabilístico.

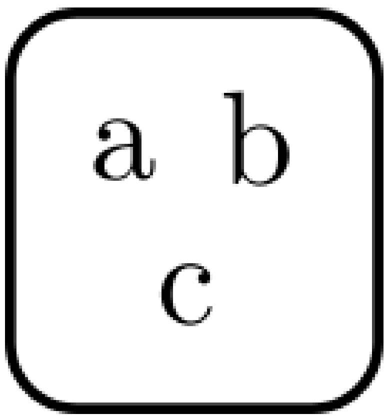
En efecto, si nos limitamos a observar los axiomas en cada escuela, encontraremos que son válidos. Sin embargo, es fácil ir directamente al análisis de los axiomas y pasar por alto si la base sobre la que están fundados es pertinente o no en las distintas escuelas de la probabilidad.

Antes de considerar siquiera si una escuela de la probabilidad cumple con la axiomática de Kolmogorov, es necesario tener un espacio muestral Ω que sea un conjunto matemático, un espacio de eventos F que sea una σ-álgebra y una probabilidad P que sea una función no negativa con dominio en F.

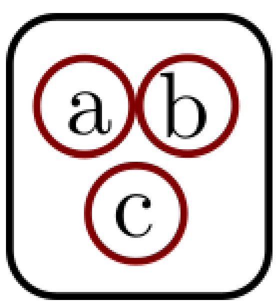
Solo en el caso de cumplir lo anterior, tiene sentido comparar a una escuela de la probabilidad con la axiomática de Kolmogorov.

2.4. ¿Qué es una σ-álgebra?

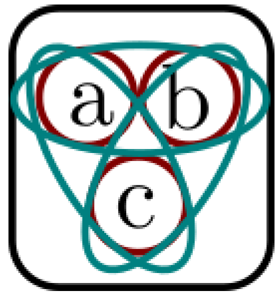
Empecemos por entender el concepto de σ-álgebra. El conjunto Ω se refiere a cualquier conjunto matemático, vamos a poner por ejemplo el conjunto $\{a, b, c\}$. La σ-álgebra F se refiere a una familia de subconjuntos de nuestro conjunto Ω que sea cerrado ante intersecciones y uniones. Por ejemplo, la F más trivial que cumple ser σ-álgebra es la que tiene al conjunto vacío y el conjunto Ω. La unión entre Ω y el vacío es Ω y la intersección entre éstos es el vacío. Esto demuestra la cerradura bajo la operación intersección y unión.



Pensemos ahora en $F = \{\{a\}, \{b\}, \{c\}\}$ que ponemos de manera gráfica en la figura 2.4. Esta familia de conjunto no es una σ-álgebra pues podemos unir $\{a\}$ con $\{b\}$ obteniendo $\{a, b\}$ que no es elemento de F. Esto lo podemos ver gráficamente en la siguiente figura.



Para lograr que la F propuesta sea una σ-álgebra se deben agregar los conjuntos $\{a, b\}, \{a, c\}, \{b, c\}$ y $\{a, b, c\}$. A toda la familia de subconjuntos se le llama la *potencia* del conjunto y siempre cumple con ser una σ-álgebra. Esta la podemos encontrar en la siguiente figura:



Para relacionarlo con el uso común en probabilidad, pongamos otro ejemplo. Supongamos dos variables aleatorias X y Y que solo pueden como valor 0 y 1. Hagamos una medición de ambas variables y supongamos que ambas tienen como resultado 0. Acordaremos en representar esto con $(0, 0)$. Si la medición resulta en 1 para X y 0 para Y lo representaremos con $(1, 0)$. Repetiremos de manera análoga para cada una de las cuatro posibilidades.

El párrafo anterior ya tiene implícita toda la estructura de la que hemos venido hablando. En este ejemplo el espacio muestral es el conjunto matemático $\Omega = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Es decir el espacio de todos los posibles resultados. Pasemos al espacio muestral F. Este no está especificado, sin embargo, casi siempre se asume que consiste en todos los posibles subconjuntos de Ω.

$$\begin{aligned} F = \Big\{ & \{(0,0)\}, \{(0,1)\}, \{(1,0)\}, \{(1,1)\}, \\ & \{(0,0),(0,1)\}, \{(0,0),(1,0)\}, \{(0,0),(1,1)\}, ... \\ & \{(0,0),(0,1),(1,0)\}, \{(0,0),(0,1),(1,1)\}, ... \\ & \{(0,0),(0,1),(1,0),(1,1)\}, \Big\} \end{aligned} \tag{2.1}$$

Este espacio se denomina de eventos pues a cada elemento se le asocia un evento en el sentido cotidiano de la palabra. Por ejemplo, el evento $X = 1$ se le asocia el conjunto $\{(1, 0), (1, 1)\}$, esto es todas las parejas que tienen en la primera entrada un uno. Para el evento $Y = 1$ asociamos el conjunto $\{(0, 1), (1, 1)\}$. Si pensamos en el evento $X = 1 \wedge Y = 1$ tendremos el conjunto $\{(1, 1)\}$. Aquí podemos ver que se cumple la siguiente propiedad:

$$\begin{aligned} (X = 1 \wedge Y = 1) = \{(1,1)\} = \\ \{(1,0),(1,1)\} \bigcap \{(0,1),(1,1)\} = (X = 1) \bigcap (Y = 1) \end{aligned} \tag{2.2}$$

Gráficamente podemos pensar que las conjunciones entre eventos equivalen a intersecciones entre conjuntos.

Para la disyunción entre eventos, tendremos una relación análoga pero utilizando la unión de conjuntos en vez de la intersección. Por ejemplo:

$$(X = 1 \vee Y = 1) = \{(1,0),(1,1)\} \bigcup \{(0,1),(1,1)\} = \\ \{(0,1),(1,0),(1,1)\} = (X = 1) \bigcup (Y = 1) \tag{2.3}$$

Notemos de nuevo que el conjunto consiste en todas los pares ordenados donde alguna de las dos condiciones se verifica. Además de que este conjunto también se encuentra en F.

Incluso podemos pensar en el evento $X = 1 \wedge X = 0$ que parece contradictorio; la variable X no puede tener ambos resultados en el mismo evento. Dentro de nuestro espacio podemos ver que:

$$(X = 0 \wedge X = 1) = (X = 0) \bigcap (X = 1) = \\ \{(0,0),(0,1)\} \bigcap \{(1,0),(1,1)\} = \emptyset \tag{2.4}$$

Es decir, para estas posibilidades contradictorias le asociaremos el conjunto vacío.

Una vez que ya explicamos la estructura del espacio de eventos pasemos a hablar sobre la relación que tiene con la función de probabilidad.

Primero, recordemos que la probabilidad se evalúa en el espacio de eventos. Por un ejemplo, la probabilidad que comúnmente se escribe $P(X = 0, Y = 0)$ tiene su formalización matemática como $P(\{(0,0)\})$. Por otro lado, la probabilidad que normalmente representamos con $P(X = 0)$ corresponde a la expresión $P(\{(0,0),(0,1)\})$ y así con todos los posibles eventos

Como mencionamos, *es un requisito para la función de probabilidad, P, estar definida para cada uno de los elementos del espacio de eventos*. En el caso anterior esto se aterriza de la siguiente manera: si las probabilidades $P(X = 1)$ y $P(Y = 1)$ están definidas, la probabilidad $P(X = 1 \wedge Y = 1))$ está necesariamente definida. Esto es una consecuencia de la exigencia sobre el espacio de eventos y la función de probabilidad sin necesidad de considerar otros elementos.

Volviendo a la divergencia entre los axiomas de Kolmogorov y las escuelas de la probabilidad; algunas escuelas niegan que los eventos realmente siguen la estructura de σ-álgebra. Es decir, no es cerrado ante uniones e intersecciones.

Por otro lado, también es posible aceptar que los eventos sí tienen dicha estructura; sin embargo, todavía es posible negar que la probabilidad está definida para todos los elementos de ésta. Es decir, el *dominio de la función* no es todo el espacio de eventos, solo un subconjunto de estos. Como consecuencia hay eventos a los que no se les puede asociar una probabilidad.

Pasemos ahora a hablar de las distintas escuelas de la probabilidad para ver cómo difieren de los axiomas antes discutidos.

Capítulo 3

Cuatro probabilidades

Vamos a discutir los cuatro puntos cardinales de la probabilidad según la bibliografía actual. Agregamos una que es significativa para el desarrollo histórico de la probabilidad. Estos ejemplos no aspiran a ser una lista completa, sino solo unos ejemplos significativos. Actualmente hay algunos ejemplos de probabilidades que realmente rompen los esquemas que vamos a proponer, pero sería muy difícil hacer una lista más completa.

Incluso estas probabilidades de las que hablaremos no son absolutas. En cada escuela de pensamiento se pueden encontrar autores que difieren en puntos cruciales. Por ejemplo, el frecuentismo de von Mises contra el de Reichenberg. Pero de nuevo, requeriría un espacio mucho más amplio de lo que este trabajo puede dar.

Finalmente, me disculpo por no incluir las probabilidades intersubjetivas que tienen apenas unos 30 o 40 años de vida, pero que realmente abren posibilidades no antes vistas. Su complejidad rebasaría este trabajo, pero no podía dejar de decir algunas palabras sobre ellas pues rebasan a sus antecesoras en capacidad y abren la puerta a la especulación de lo que se podría hacer con ellas en terrenos de la física.

3.1. La teoría clásica de la probabilidad

La visión clásica de la probabilidad se refiere a la que se conformó en los siglos XVI y XVII para ser la predominante hasta el siglo XIX e incluso seguir vigente entre algunos científicos a principios del siglo XX [36, p. 18]. Su gran difusión se debe a textos clásicos como el "Ensayo filosófico sobre las probabilidades" de Laplace que estaba dirigido al "lector educado"[18, p. III]. Esta visión está directamente relacionada con el pensamiento de la ilustración e intentaba ser “un modelo matemático para este tipo de racionalidad rutinaria”[19, p. 50].

Para la visión de Laplace, muy influida por el éxito de la mecánica newtoniana y dentro de la ilustración, el mundo está regido por leyes naturales que describen perfectamente el movimiento de las cosas en cualquier momento del tiempo. Así como la mecánica lleva a ecuaciones diferenciales cuando se aplica a un caso en particular y estas requieren de condiciones iniciales para determinar las constantes de la solución, en principio se podría describir el curso de todo si tenemos el conocimiento de las leyes y de las condiciones iniciales. Sin embargo, no siempre están siempre a nuestro alcance lo que nos pone una limitante.

Una persona aplica la probabilidad debido a que no tiene los elementos antes mencionados, pues en el fondo lanzar moneda está “determinada” a tener un valor que no conocemos. Laplace lo pone de la siguiente manera: En esta visión, la probabilidad depende directamente de quien que la usa y de lo que sabe sobre una situación dada. Un ejemplo sencillo es un dado tramposo que no tiene el número 6, cuando preguntamos a las personas cuál es la probabilidad de sacar 6 dirán que es un sexto, pero al enterarse de la condición del dado dirán que es cero.

Otro elemento importante es la asignación de probabilidades. En la probabilidad clásica se asignan probabilidades iguales a todos los eventos que nos parezcan equivalentes. “La teoría de las posibilidades”, dice Laplace, “consiste en reducir todos los eventos del mismo tipo a un número de casos igualmente posibles”[51, pp. 6-7]. Ésta es la regla de asignación que ya hemos tocado antes y es familiar para la mayoría de las personas. Como habíamos mencionado, esta regla podía dar valores distintos de probabilidad y consideramos esto un problema. Sin embargo, para la probabilidad clásica no hay ningún problema pues ésta es relativa a la ignorancia de quien la aplica.

Ya habíamos tocado el ejemplo de lanzar dos dados, por un camino asignamos al número 12 la probabilidad $1/36$ y por otro $1/11$. Cuando revisamos el argumento de $1/11$ nos damos cuenta que no hace uso del hecho que los números vengan de la suma de dos dados, es decir, el mismo argumento se pudo usar si los número vinieran de la lotería. Por tanto una persona que desconoce la lista completa de las posibilidades de números podría asignar esta probabilidad. Por otro lado, para usar el primer argumento sí es necesario saber que los números saldrán de la suma de dos dados. En definitiva, estas asignaciones dependen de la ignorancia. También es posible justificar por qué el resultado $1/36$ sí coincide con los resultados del juego cuando se repite muchas veces pues este tiene “información extra” que el otro carece, poniéndolo así en ventaja.

Finalmente, vamos a abordar otro problema aparente que tiene esta aproximación desde la visión cotidiana que discutimos antes. La forma de asignar probabilidades parece no parece dar lugar a los resultados de muchas mediciones, por ejemplo, si tenemos una moneda que casi siempre sale "águila", la regla aún nos indica asociarle una probabilidad un medio, lo que no parece adecuado. Aquí, de nuevo debemos tomar en cuenta el conocimiento de quien asocia las probabilidades. El capítulo VII de su tratado, Laplace considera esta situación y afirma: "... la probabilidad de tirar cara en la primera tirada será siempre 1/2; debido a nuestra ignorancia de cuál cara está favorecida..."[51, p. 56] De nuevo es necesario considerar la ignorancia de la condición de la moneda, lo que equivale a considerar que una cara es más pesada que la otra. No obstante, una persona que sí conoce tal condición, no consideraría los casos como equivalentes y por tanto no asignaría el valor 1/2.

Solo nombraremos estos elementos de la probabilidad clásica. Naturalmente, esta visión es más amplia y hay otros elementos que pueden ser relevantes, sin embargo nos restringimos a estos por ser los necesarios para nuestra discusión.

3.2. La teoría logicista de la probabilidad

La visión logicista de la probabilidad tiene su origen a finales del siglo XIX y principios del XX en Cambridge. Esta visión fue desarrollada por personas que tuvieron gran influencia en el pensamiento del siglo XX. Nombres como Russell, Moore, Keynes figuran en los círculos relacionados a esta visión. Después también se desarrolló por algunos integrantes del círculo de Vienna. Por ejemplo, Carnap escribió ampliamente al respecto [15].

Esta visión tiene muchas similitudes con la visión clásica, pero se reinterpretan varios elementos en términos lógicos. Empecemos con el concepto de probabilidad, al respecto Keynes dice: "Los términos cierto y probable describen los varios grados de creencia racional acerca de una proposición que diferentes cantidades de conocimiento nos autoriza a considerar"[43, cap.1 § 2]. Como vemos, la probabilidad se le va a asociar a proposiciones y conocimiento, en contraste con una vía usual de asociar la probabilidad a eventos. De manera análoga a la visión clásica, el valor de probabilidad no depende solo de la proposición, también del conocimiento previo o dado. Aquí ese conocimiento se identificará con las premisas, de igual manera que en un teorema. Veamos como lo pone Keynes: "Sean nuestras premisas cualquier conjunto de proposiciones h, y nuestra conclusión cualquier conjunto de proposiciones a, entonces, si un conocimiento de h justifica una creencia racional en a de grado α, decimos que hay una relación de probabilidad de grado α entre a y h."[43, cap. 1 § 3] Aquí vemos como el "conocimiento previo" se identifica con un conjunto de premisas. Gilles, al hablar de Keynes, caracteriza esto como una "implicación parcial" [36, p. 30] . Es decir, así como un conjunto de premisas pueden implicar un teorema, aquí una serie de premisas implican parcialmente una proposición. No la aseguran, pero sí nos llevan a una creencia racional en tal proposición.

Nos gustaría abundar sobre la idea de que la probabilidad depende de las premisas. Bajo esta visión, es imposible asociar un valor de probabilidad a una proposición si no se tienen premisas. De la misma manera que una proposición no puede demostrarse si no se señalan las premisas de las que se parte, con la excepción de las tautologías. A primera vista, podríamos pensar que esto es incorrecto y señalar uno de los muchos ejemplos donde asociamos probabilidad sin premisas, como la moneda, por ejemplo. Sin embargo, a esto la escuela logicista nos responde que siempre hay unas premisas tácitas para toda situación. Volviendo a la moneda, ya hay un conocimiento previo compartido por las personas que utilizan la moneda, a saber, que moneda tiene dos lados, que se tiene que lanzar de una cierta manera para que el tiro se considere adecuado, etc.

Vamos a discutir ahora la asignación de valores. Este punto es sin duda uno de los más complicados para esta visión pues es crítica con respecto a la asignación de probabilidades. Primero es importante mencionar que en la definición el grado α de creencia racional no es un número directamente, sino un grado al que le podemos asociar un número, de la misma manera que la longitud no es un número sino una magnitud o propiedad a la que dadas una condiciones y un sistema de medición le asociamos un número. Veremos que asociarle números a las probabilidades no es una labor trivial. Keynes lo escribe así: " Ha sido a veces supuesto que una comparación numérica entre los grados de cualquier par de probabilidades no solo es concebible sino que está en nuestras manos... Que esa comparación es teóricamente posible, seamos o no competentes en cada caso para hacer la comparación, ha sido una opinión aceptada generalmente." [43, cap. III § 1]. No aceptar que la comparación de probabilidades es siempre posible rompe fundamentalmente con una visión donde las probabilidades tienen un número asignado.

Para poder asociar un número a alguna propiedad es una condición indispensable que ésta sea ordenada. Es decir, que al menos en principio podamos comparar dos propiedades y decir cuál es mayor o si son iguales [25, pp. 25-28]. Negar esto hace imposible el ordenamiento y por tanto la asignación de números. Como ejemplo de esta situación Keynes escribe: "Considere tres conjuntos de experimentos, cada uno con el objetivo de establecer una generalización. El primer conjunto es más numeroso; en el segundo el conjunto de condiciones irrelevantes ha sido variadas con más cuidado; en el tercer caso la generalización en cuestión es más amplia en alcance que en las otras. ¿Cuál de estas generalizaciones es bajo tal evidencia la más probable?" [43, cap. III § 8]. Él afirma que no es posible comparar estos casos, además que no hay forma de resolver siquiera si son iguales o no sus probabilidades.

Con lo anterior no se refiere a que no conozcamos un método para comparar o que éste es tan complicado que no es posible llevarlo a la práctica sino a que no existe ninguno: "Ningún método de cálculo, aún muy impracticable, ha sido sugerido. Tampoco tenemos ninguna indicación *prima facie* de la existencia de una unidad común a la cual las magnitudes de todas las probabilidades son referibles naturalmente." [43, cap. III § 9] Es decir, ni siquiera en principio se tiene una forma de realizar la asignación. Este punto tan delicado está ampliamente discutido en el capítulo mencionado donde, además, se dan ejemplos cotidianos como las aseguradoras y algunos casos legales que revelan la imposibilidad de asignar valores numéricos o definir siquiera un orden total en esas situaciones. Aunque no lo menciona, podemos apreciar gran similitud con las ideas de Helmholtz y Campbell sobre la medición y construcción de escalas cuantitativas; lo que actualmente se conoce como la teoría de la medición o de la metrización.

Si bien no es posible comparar todas las probabilidades, sabemos que en algunos casos sí la podemos comparar. Es decir, no es posible poner un orden total en las probabilidades, pero sí un orden parcial que nos permitirá asignar números para una parte de las situaciones. Como ayuda gráfica, Keynes nos da el diagrama de la figura 3.1 que representa las posibles asignaciones de probabilidad [43, cap. III § 20]. La recta en la parte inferior representa el intervalo cero uno en el cual la probabilidad puede tomar valores numéricos. Como primera característica vemos que todas las probabilidades se pueden comparar con la imposibilidad O y la certidumbre 1. Por un lado tenemos la probabilidad A a la que sí se le puede asignar un número, por eso se encuentra sobre la recta. Por otro lado, las probabilidades U, V, W, X, Y, Z; representan algunas para las que no es posible hacer tal asignación. Entre estos casos tenemos V que sí podemos comparar con A y vemos que es menor, pero no es posible asignarle un número. V también se puede comparar con W y sabemos que es menor, para Z tenemos la misma situación. Sin embargo, V y Z no se pueden comparar entre sí. Una situación análoga la tienen X y Y pues son mayores que W pero incomparables entre sí. Finalmente, tenemos que U no es comparable con ninguno de los otros casos. Nos parece que este diagrama pone de relieve la complejidad de la probabilidad en esta visión

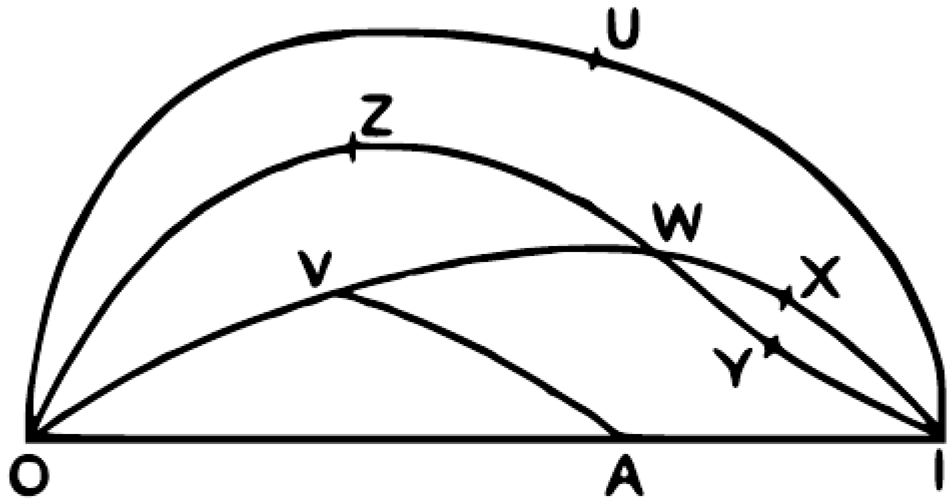


Figura 3.1: Esquema que representa las distintas probabilidades en la visión de Keynes

en particular cuando se contrasta con cualquier otra aproximación que sólo se tiene como posibilidades las que están en el intervalo OI.

¿Cómo se asocian los valores numéricos en esta visión? A través de una regla, análoga a la de la visión clásica, llamada por Keynes el principio de indiferencia. Pasemos a explicar los elementos para establecerlo. Para asociar probabilidades iguales en la visión clásica se toma la colección de eventos equivalentes y se les asocia la misma probabilidad a todos. Sin embargo, esta visión no nos proveía de un criterio estricto para establecer qué casos deben ser considerados equivalentes. En la escuela logicista sí tenemos este criterio y, dado que esta probabilidad se asocia a proposiciones, exigirá que las proposiciones equivalentes tengan la misma "forma"; esto es, que se puedan escribir como $\phi(a)$, $\phi(b)$, etc. donde ϕ es una proposición de la lógica formal.[1] Por ejemplo, un par de proposiciones equivalentes son "el libro es rojo" y "el libro es azul", ambas se pueden escribir como $\phi(rojo)$ y $\phi(azul)$. Por otro lado, las proposiciones "el libro es rojo" y "el libro no es rojo" no son equivalentes porque "el libro no es rojo" no se puede poner en la forma

[1] Aunque Keynes no abunda en lo anterior, hay una serie de condiciones necesarias para tal proposición $\phi(x)$.

$\phi(x)$, pues $\phi(norojo)$, es decir "el libro es no rojo", no es una oración bien formulada.

Este principio tiene otra regla de aplicación que no encontramos en la visión clásica explícitamente: "...es una condición necesaria para la aplicación del principio que éstas [las proposiciones] sean, relativamente a la evidencia, indivisibles alternativas de la forma $\phi(x)$". Volviendo al ejemplo de los libros, la proposición ϕ(rojo y azul) puede ser dividida en $\phi(rojo) \wedge \phi(azul)$. Algunas de las paradojas más comunes en la asignación de probabilidades tienen solución con este principio. Veamos una de ellas[36, pp. 37-46]. De tres libros, uno rojo, uno azul y uno verde; escogemos uno al azar ¿cuál es la probabilidad de sacar el libro rojo? Uno de los caminos es tomar las proposiciones "el libro es rojo" y a "el libro no es rojo" y darles la misma probabilidad, 1/2 en este caso. Como mencionamos antes, estas proposiciones no son equivalentes, pues no tienen la forma $\phi(x)$, por lo que el principio de indiferencia no se puede aplicar. Por otro lado, podríamos intentar el par de proposiciones "el libro es rojo" y "el libro es azul o verde" para usar el principio de indiferencia y darles probabilidad 1/2. De nuevo, no cumplimos con las condiciones, en este caso la de indivisibilidad. Pues "el libro es azul o verde", como ya mencionamos, sí es divisible. Así solo nos queda la opción tomar los enunciados "el libro es rojo", "el libro es azul" y "el libro es verde"; que sí cumplen con la equivalencia y la indivisibilidad, y asociarles 1/3 de probabilidad a cada uno.

Gilles discute a profundidad este principio aplicado a algunas de las paradojas más famosas. En particular, las asignaciones de probabilidad a intervalos continuos que son importantes para nuestro caso y para la probabilidad geométrica. Keynes, también discute esto en relación a la probabilidad geométrica. Para él no es posible formular directamente las probabilidades en un intervalo pues siempre son divisibles tal como mencionamos en el párrafo anterior. En concreto "x está en $[0,1]$" siempre se puede dividir en "x está en $[0,1/2]$" o "x está en $[1/2,1]$". Sin embargo, propone otras formas indirectas de representarlo que abordaremos más adelante.

3.3. La escuela frecuentista

La escuela frecuentista comenzó a mediados del siglo XIX en Inglaterra con los trabajos de Venn y continuó hasta la primera mitad del siglo pasado donde se desarrolló más extensamente. En particular estamos interesados en Richard von Mises, una de las figuras más representativas de esta aproximación, al menos en relación a la física y en particular a la mecánica cuántica.

Él tuvo una carrera muy prolífica en distintas ramas de la física y se dedicó durante muchos años al desarrollo de los fundamentos de la probabilidad en una mirada muy a tono con el círculo de Vienna. Aunque publicó diversos artículos al respecto, su trabajo de referencia es *Probability, Statistics and Truth* de 1928. Nos parece relevante su aproximación pues es a von Mises a quien se hace referencia en las pocas veces que los autores originales de la mecánica cuántica hablan de probabilidad[33].

Mientras que las otras miradas buscan incorporar los distintos usos de la probabilidad en varias situaciones, esta mirada tiene como principal objetivo capturar el carácter empírico de la probabilidad. El hecho de que al lanzar una moneda muchas veces las frecuencias relativas se aproximan a un valor definido se considera el fundamento de toda la teoría. Von Mises lo dice explícitamente: *It is essential for the theory of probability that experience has shown that in the game of dice, as in all the other mass phenomena which we have mentioned, the relative frequencies of certain attributes become more and more stable as the number of observations is increased.* [55, pág. 12]. Ese valor al que aceptamos que las frecuencias relativas se estabilizarían al seguir realizando observaciones será el que definamos como la probabilidad.

Definida así, la teoría de la probabilidad se convierte en una teoría "física" que está basada en experimentos, donde la probabilidad se mide. Von Mises asegura que la probabilidad es una variable física: *The probability of a 6 is a physical property of a given die and is a property analogous to its mass, specific heat, or electrical resistance. …The theory of probability is only concerned with relations existing between physical quantities of this kind.* [55, pág. 14]. Al ser un propiedad de un sistema la probabilidad se mide para cada caso. Por tanto, no es correcto hablar de "la probabilidad de un dado" y usar los principios de asignación de probabilidades que hablamos antes, más bien tiene sentido hablar de la "probabilidad de ese dado" o "aquel dado". De esta manera podemos justificar nuestra creencia de que hay dados "justos" e "injustos", los justos son los que cumplen con nuestro ideal de dado, cosa que solo podemos verificar midiendo.

Además de dados y otro tipo de eventos que se repiten, el concepto de probabilidad se aplica a "eventos en masa", eventos que se repiten de manera análoga para una población grande. Por ejemplo, tenemos las edades de las personas de una ciudad. Una pregunta común sería cuál es la probabilidad de que alguien de tal ciudad tenga una edad de 40 años. Al tomar personas al azar y preguntar sus edades tenemos una situación similar a la del dado que se tira una y otra vez. Von Mises también insiste en los ejemplos de la física. Por ejemplo la velocidad de las partículas de un gas, donde se puede plantear la misma cuestión en relación a las velocidades.

En definitiva, la probabilidad estará bien definida siempre y cuando se especifiquen unas condiciones respecto de las cuales se repetirá un evento o una población grande a la cual se le aplicaría un experimento. Estos dos casos quedan englobados en lo que von Mises llama el colectivo y define así: *We will say that a collective is a mass phenomenon or a repetitive event, or, simply, a long sequence of observations for which there are sufficient reasons to believe that the relative frecuency of the observed attribute would tend to a fixed limit if the observations were indefinitely continued. This limit will be called the probability of the attribute considered within the given collective.* [55, pág. 15]. Aquí podemos observar que la probabilidad dependerá completamente del colectivo que se use para definirla, no solo eso, la falta de un colectivo prohibirá la existencia de la probabilidad: *'The probability of winning a battle' , for instance, has no place in our theory of probability, because we cannot think of a collective to which it belongs* [55, pág. 15]. Con respecto a lo anterior, Gilles afirma que aunque esta probabilidad es objetiva, comparte con las probabilidades epistémicas que será siempre condicional. Mientras que las primeras están condicionadas al conocimiento o hipótesis con que se aplica la probabilidad, la segunda está condicionada a la elección del colectivo.

Hay otra condición que von Mises impone a una secuencia para que sea un colectivo y pueda dar lugar a una probabilidad. Hay muchas secuencias de mediciones donde las frecuencias relativas se estabilizan que sin embargo partes de ellas no siguen las mismas frecuencias. Por ejemplo, una moneda muy extraña que tiene como resultado Águila, después Sol, luego Águila, Sol y así sucesivamente. Aunque es inmediato que la frecuencia de Soles se estabilizará rápidamente en el valor $1/2$, la subsucesión que solo toma los lanzamientos pares tendrá frecuencias relativas distintas a las de la secuencia original, pues serán únicamente Soles. Cuando una secuencia tiene alguna relación entre el número de la muestra y el resultado podremos escoger una subsucesión que rompa con las frecuencias relativa originales. De esta manera terminamos con dos condiciones para el colectivo: que tengan frecuencias relativas que se estabilicen para muchos eventos o muestras y que la sucesión de resultados no tengan patrones que nos permitan cambiar las frecuencias relativas para una subsucesión.

Las dificultades que se enfrentan para explicar la consistencia de las frecuencias relativas aquí quedan eliminadas pues están en el centro de la visión. Sin embargo, otras cuestiones surgen en esta misma relación. A pesar de que las probabilidades se definen como el límite de estas frecuencias es de notar que en cualquier caso real, el número de observaciones, repeticiones o experimentos es finito. Esto representa un problema en particular desde el punto de vista del empirismo y otras visiones afines, con las que von Mises tenía mucha relación. Von Mises es consciente de este problema, pero ve a las secuencias finitas como análogas a los puntos materiales de la mecánica clásica, pues en ambos casos se entienden como abstracciones matemáticas que utiliza la teoría y solo sirven como punto de referencia para su desarrollo. Finalmente, aunque estas idealizaciones no se fundamenten adecuadamente o presenten problemas, se soportan debido al éxito de la teoría para relacionarse con los datos empíricos y en esta visión de la probabilidad como una ciencia empírica, su éxito es suficiente como para fundamentar el uso de secuencias infinitas.

3.4. Las propensiones o propensidades

También dentro de las probabilidades objetivas tenemos las propensiones de Popper que han dado lugar a una gran variedad de teorías que se acercan de alguna u otra manera a la idea original. Él estaba de acuerdo con la interpretación frecuentista de la probabilidad, sin embargo, la imposibilidad de asociar probabilidades objetivas a eventos individuales le parece un gran inconveniente que debe resolverse. Para Popper este punto es de particular interés pues se relaciona con su interpretación de la mecánica cuántica [36, 62, pág.115].

La idea original para asociar una probabilidad objetiva a un evento individual consiste en asociar al evento individual con un colectivo y darle la misma probabilidad que éste. Cuando nos referimos a un colectivo es común asociarlo con condiciones en las que se podría realizar. Por ejemplo, un colectivo se puede definir como los eventos que vienen de tirar una moneda dada en unas condiciones definidas, lo que tiene como ventaja que elimina la necesidad de especificar cada uno de los elementos que compondrán el colectivo. Esto sugiere que se puede asociar la probabilidad con las condiciones en las que un evento sucede y no necesariamente con el colectivo.

Pongámoslo así, cuando lanzamos una moneda, las condiciones con las que la lanzamos son propensas a que el resultado sea águila. En ese sentido se habla de propensión, disposición o tendencia, es decir qué tan propensas son nuestras condiciones a producir resultados cuyas frecuencias relativas coincidan con el valor de la probabilidad.

En general esto permite dar probabilidades objetivas a eventos individuales. Estrictamente no se asocian al evento directamente, más bien tendremos probabilidades objetivas que se asocian a condiciones y naturalmente le podemos asociar la probabilidad a un solo evento asignándole la probabilidad de las condiciones en que se realizó sin necesidad de repetir el evento muchas veces.

El principal problema de esta aproximación es la definición de estas condiciones que de manera similar se encuentra en la probabilidad frecuentista. Debemos elegir un conjunto de condiciones particular que describa al evento por encima de todos los posibles conjuntos de condiciones que lo describan. Para esto se tiene el *principio de la clase de referencia más pequeña*. Éste indica que para describir un evento se tienen que utilizar las condiciones más específicas posibles.

3.5. La escuela subjetivista

La segunda escuela de pensamiento que vamos a revisar es la subjetivista. Ésta es una aproximación que se califica de epistémica o epistemológica pues la probabilidad dependerá del conocimiento de quien la aplica. Mientras que para los logicistas la probabilidad es una medida de creencia racional, los subjetivistas negarán la racionalidad y se refieren a las creencias, junto con la probabilidad, a un ámbito más amplio.

La visión subjetivista tiene dos autores principales: Frank P. Ramsey y Bruno De Finetti. Ambos desarrollaron sus ideas a principios del siglo XX de manera independiente. Por un lado, Ramsey pertenecía a las élites en Cambridge y sus trabajos en buena medida respondían a la escuela logicista y directamente a Keynes, por otro lado, de Finetti se encontraba en Italia donde no había un dominio tan fuerte de alguna escuela. Aquí nos centraremos en de Finetti pues Ramsey vivió pocos años luego de publicar sus ideas; desgraciadamente tuvo una muerte temprana a los 26 años. Sin embargo, sus visiones no son equivalentes y hay diferencias fundamentales en su aproximación que han sido tratadas ampliamente en [32].

La escuela subjetivista acepta que las probabilidades dependen del conocimiento de quien la aplica, pero a diferencia de los logicistas, no se limitan a una relación racional entre este conocimiento y la probabilidad. Ya no se habla de una creencia racional sino solo de creencias en general. De Finetti lo expresa así en uno de sus textos más famosos "... only subjective probabilities exist- i.e. the degree of belief in the occurrence of an event attributed by a given person at a given instant and with a given set of information" [20]. Así, no hay tal cosa como "la probabilidad" sino solo hay probabilidades respecto a las personas que la aplican, es decir, "la probabilidad según Juan de que tal cosa pase", "la probabilidad según Pablo de que ...", etc.

Esta actitud parece radical cuando se confronta con la visión cotidiana de la probabilidad, después de todo aseverar que las probabilidades dependen de lo que las personas creen es equivalente a negar que la probabilidad tiene algo de objetivo. Esta postura es defendida por de Finetti en todos sus textos y desde distintos ángulos. Por ejemplo, en Probabilismo, uno de sus artículos más representativos, podemos encontrar: "By denying any objective value to probability I mean to say that, however an individual evaluates the probability of a particular event, no experience can prove him right, or wrong; nor, in general, could any conceivable criterion give any objective sense to the distinction one would like to draw, here, between right and wrong."[21, § 5]. De Finetti no está en contra de la objetividad directamente sino que no encuentra un camino viable para ésta.

Volvamos por ahora a las probabilidades subjetivas. Debido a que la probabilidad depende de la persona que la aplica, podemos tener distintos valores de la probabilidad para la misma situación e incluso para el mismo conocimiento de parte de la persona, a diferencia de las probabilidades logicistas. En el ejemplo de la suma de dos dados que tratamos antes, la probabilidad de que salga 12 puede tomar cualquier valor, 0, 1/2, 1/29, 1/300, etc. A primera vista podríamos pensar que las probabilidades que asocian las personas a los eventos podrían ser falibles, tramposas, inconsistentes, por ejemplo, a una persona le preguntamos las probabilidades en un volado y nos responde 50 % para águila y 75 % para sol. Esto da una imagen negativa de esta visión, sin embargo, debemos ver con cuidado la definición de de Finetti, pues él asegura que la probabilidad es el "grado de creencia" que tiene una persona sobre un evento. Este grado no necesariamente es conocido por las personas y no necesariamente están en posibilidad de decir un número que represente su creencia. Se necesita un método para medir estos grados de creencia que sea fiable.

De Finetti está muy en sintonía con el operacionalismo, pues para él los problemas básicos de la fundamentación de la probabilidad son "qué es la probabilidad" y "cómo se mide". Si se logra responder la primera y establecer un método coherente para medirla se tiene buena parte del camino resuelto [36, p. 58]. El método que propone la escuela subjetivista, tanto del lado de Ramsey como de de Finetti, se encuentra en un juego de apuestas que "revele" las creencias de la persona. La propuesta es la siguiente, dado un evento y una persona, queremos saber el grado de creencia de ésta en que el evento va a pasar, es decir la probabilidad. La persona tiene que elegir un número q, que llamaremos el coeficiente de apuesta, y nosotros escogeremos una cantidad de dinero S. Ella nos pagará la cantidad qS y nosotros le entregaremos la cantidad S cada vez que el evento en cuestión ocurra. El punto fundamental de este esquema es que nuestra elección S puede ser positiva o negativa.

Vamos con un ejemplo para aterrizar este esquema que a primera vista es complicado. Apostaremos sobre el evento “que salga sol en un volado”, la persona escoge el coeficiente $q = 1$, luego nosotros escogemos como S la cantidad de 1 peso. Entonces, cada vez que sale sol tenemos que dar un peso y recibimos un peso. Al final todos se quedan con la misma cantidad de dinero. Visto lo anterior, la persona puede elegir el coeficiente $q = 0$ y nosotros $S = 1$ peso, así cada vez que sale águila nosotros damos un peso pero no recibimos dinero. Esto representa pérdidas para nosotros. Sin embargo, también podemos poner la situación donde la persona escoge $q = 0$ y nosotros $S = -1$ y cada vez que sale sol nosotros recibimos 1 peso y no tenemos que dar nada a cambio. En el medio habrá un valor para q donde para muchos eventos no habrá ganancias para nadie. Aquí reluce la importancia de que la persona tiene que elegir primero el valor q antes de que nosotros digamos si S será positivo o negativo. El coeficiente q, asegura de Finetti, será una buena representación del grado de creencia en ese evento y por tanto podemos igualarlo a la probabilidad.

Como ya dijimos, para un operacionalista, esta visión ya cumplió con las principales cuestiones que necesita responder una visión de la probabilidad. Sin embargo, todavía es necesario esclarecer cómo podemos asegurar que los coeficientes que uno obtiene con esta forma de medir, siempre resultan en números que se les puede llamar probabilidades. Bien podríamos tener que los números que resultan de este método le terminan por dar 50 % de probabilidad a ’águila’ y 75 % a ’Sol’ en una tirada de moneda, de manera similar al ejemplo que expusimos antes. Para atacar este punto vamos a pedir una hipótesis extra: las probabilidades que la persona asocia a un conjunto de eventos q_1, q_2, etc., tienen que ser tales que no sea posible construir una estrategia donde la persona siempre pierda sin importar el resultado del evento, dicho de otra manera, no es posible escoger probabilidades que permitan perder siempre. A un conjunto de probabilidades de este tipo se le llama coherente. Este nuevo requerimiento nos parece razonable dado que ya se ha planteado todo el esquema de las probabilidades dentro de las apuestas y que se tiene implícitamente que las personas actuarán de la manera que obtengan más dinero. Uno de los resultados más importantes para los subjetivistas es que esta hipótesis es suficiente para asegurar que las probabilidades asignadas cumplirán con los axiomas de la probabilidad de Kolmogorov, los que ya enunciamos antes. A este resultado se le conoce como el teorema de Ramsey-De Finetti cuyo enunciado es: Un conjunto de apuestas es coherente si y solo si satisface los axiomas de la probabilidad [36, p. 59]. Esto termina por asegurar que la interpretación y el método de medición corresponden efectivamente con la probabilidad.

Por último, pasemos a la relación que tiene esta visión de la probabilidad respecto de las frecuencias relativas. Este punto es importante para las visiones epistémicas pues deben justificar cómo es que la probabilidad depende del conocimiento y al mismo tiempo está de acuerdo con las frecuencias relativas para muchos eventos. La idea general es que una persona que haya presenciado un gran número de eventos modificará sus creencias para coincidir con la frecuencia relativa. Dentro del subjetivismo este argumento se basa en la intercambiabilidad de los eventos.[36, p. 71]. Supongamos que tenemos una serie de lanzamientos de moneda, decimos que los coeficientes de apuesta q son intercambiables si todos los casos donde sale águila el mismo número de veces tienen la misma probabilidad, es decir, águila-sol-sol, sol-águila-sol y sol-sol-águila tienen el mismo coeficiente o probabilidad.

Ahora pasemos a la relación con las frecuencias relativas y pongamos un ejemplo como ayuda. Dados una serie de n lanzamientos de moneda queremos calcular la probabilidad del siguiente, $n+1$ sea 'Sol', condicionado a que ya tenemos los resultados de los primeros n. Esta probabilidad la podemos expresar de la siguiente manera:

$$P(E_{n+1} = S | E_n, E_{n-1}, \dots, E_1) = \frac{P(E_{n+1} = A, E_n, E_{n-1}, \dots, E_1)}{P(E_n, E_{n-1}, \dots, E_1)}$$

como simple aplicación de la regla de Bayes.

Primero, llamaremos a la probabilidad de obtener r veces 'Sol' en n intentos sin importar el orden ω_n^r. Por una parte, ésta nos permite calcular la probabilidad de obtener la secuencia particular de resultados que obtuvimos, es decir, $P(E_n, E_{n-1}, \dots, E_1)$. La hipótesis de intercambiabilidad nos conduce a que todos los casos con n veces sol tienen la misma probabilidad, por tanto

$$P(E_n, E_{n-1}, \dots, E_1) = \omega_n^r / C_n^r,$$

$$\begin{aligned}
P(E_{n+1} = S | E_n, E_{n-1}, \dots, E_1) &= \frac{P(E_{n+1} = S, E_n, E_{n-1}, \dots, E_1)}{P(E_n, E_{n-1}, \dots, E_1)} && (3.1) \\
&= \frac{(\omega_n / C_n^r)}{(\omega_{n+1} / C_{n+1}^{r+1})} && (3.2) \\
&= \frac{n!}{r!(n-r)!} \frac{(r+1)!(n-r)!}{(n+1)!} \frac{\omega_n}{\omega_{n+1}} && (3.3) \\
&= \frac{r+1}{n+1} \frac{\omega_n}{\omega_{n+1}} && (3.4)
\end{aligned}$$

Adhiriendo la hipótesis de que el límite de ω_n/ω_{n+1} es uno cuando n tiende a infinito, tenemos el resultado deseado: la probabilidad de una tirada, en este caso que el resultado sea 'Sol', dado el resultado de n anteriores converge a la frecuencia relativa del resultado. Así, sin importar las creencias iniciales de una persona, las probabilidades que asocia se van acercando a la frecuencia relativa que observa y distintos observadores terminarán por coincidir en el valor de probabilidad.

Para concluir este punto nos parece importante aclarar que, según de Finetti, no existen probabilidades objetivas y este argumento solo revela por qué en ocasiones nos parece que sí las hay. El punto fino de esta argumentación se encuentra en las "repeticiones de unas mismas circunstancias", que están implícitas al hablar de lanzamientos de una moneda. Este punto no le parece aceptable a de Finetti pues ve en él una subjetividad intrínseca al escoger cuales de todas las condiciones posibles se tomarán en cuenta para caracterizar una situación. Él escribe[21, § 10]:

> the concept of *trials of the same phenomenon* is arbitrary, as is in logic that of *elements of the same class* . Any two objects can be put into one class or distinct classes, but among the innumerable classes one can form there are some which are of practical interest and for which it is generally thought useful to introduce a special denomination.....But it is only a question of utility and degree; it would be vain to look for a philosophical substratum, and it would be vain to think of the concept of *trials of the same phenomenon* as something meaningful in itself.

Volveremos a este punto más adelante.

Capítulo 4

Kolmogorov Dixit

En la sección anterior, se escribieron los tres axiomas de Kolmogorov como se escriben en cualquier libro moderno de probabilidad o de estadística. En el discurso moderno, el uso de la probabilidad se acostumbra defender con estos axiomas.

Sin embargo, es de lo más curioso revisar que las escuelas de la probabilidad de las que hemos hablado no parecen estar conectadas con dichos axiomas. Incluso parece que temporalmente no tiene sentido: Kolmogorov es un matemático muy moderno, nació en 1903, ¿cómo él realizo esta axiomática si los grandes trabajos de probabilidad vienen al menos 100 años antes? Mencionamos antes que el trabajo que divulgaría la probabilidad en las élites europeas sería el libro de Laplace que data de 1812[51].

Supongamos que obviamos esto, pero cómo obviar que las élites británicas tuvieron sus trabajos más importantes de probabilidad, e.g. el libro de Keynes[43], justo en la primera década del siglo XX cuando Kolmogorov era todavía un bebé y los de Ramsey a principios de los 20's[64] cuando Kolmogorov aún no publicaba los trabajos que lo harían famoso [35, 65].

Podemos seguir con las disonancias. El libro de Kolmogorov donde se propone su famosa axiomática es de 1933[45](con la versión en ruso en 1936 y en inglés hasta los 50's), ¿cómo fue relevante si la teoría logicista, la subjetivista y hasta la teoría de von Mises ya tenían una forma acabada[36]? Es claro que estas teorías tienen choque entre sí. ¿Por qué se usan los axiomas de Kolmogorov en vez de los axiomas de alguna de estas teorías? De hecho la de von Mises tiene solo dos axiomas[55].

Incluso las aplicaciones de la probabilidad en la estadística ya habían pasado su climax para 1933. Recordemos que el trabajo más impactantes de Fisher *Statistical Methods for Research Workers* ya se habían publicado en los 20's[30, 79]. Fisher pertenecía a las élites británicas y aunque se puede debatir sobre si hay un uso de la probabilidad logicista en su trabajo, es claro que no tuvo necesidad de los axiomas de Kolmogorov en su obra.

Podríamos creer que la fundamentación de la probabilidad fue increíblemente difícil de proponer y por eso los axiomas se escribieron después de todas estas teorías. Sin embargo, nos encontramos con otro hecho muy evidente: los axiomas de Kolmogorov son muy simples ¿no lo cree? Básicamente es una teoría de la medida que se le ponen condiciones para que sume 1. En los salones de clase donde se enseña la probabilidad no es difícil encontrar alumnos que noten esto de inmediato. Parece un poco raro que con tantos trabajos sobre probabilidad a principios del siglo XX se ponga tanto énfasis en dichos axiomas que no parecen realmente resolver una cuestión nueva.

Vamos a llevar un paso más allá la pregunta, porque este tema realmente lo amerita. Si uno, esforzadamente, junta el tiempo para abrir el texto original de Kolmogorov[45] se dará cuenta de una cosa que raya en lo insólito: los axiomas de Kolmogorov son cinco, sí, cinco. Así que en un momento de completa derrota intelectual debo preguntar ¿en qué momento se nos perdieron dos axiomas? y mejor aún: ¿por qué no veo a nadie haciendo un escándalo de estas situación si estamos hablando de la herramienta matemática más relevante para la ciencia de nuestra época?

Las respuestas desgraciadamente tienen mucho menos de interesante que lo que uno podría imaginar.

4.1. El monstruo del Dr. Kolmogorov

Al entrar a una clase de probabilidad de nivel universitario, somos expuestos a un discurso de la probabilidad incoherente. Este discurso ese heredero de todas las teorías de la probabilidad que le anteceden y toma cosas de uno y de otro. Sin embargo no puede tomar de todos y seguir siendo consistente.

Hay afortunadamente una gran literatura, en español, que habla sobre la enseñanza de la probabilidad desde los niveles primarios hasta los universitarios. Ahí se puede ver que ellos son conscientes de los muchos significados de la probabilidad y se promueve que estos significados sean enseñados. Aunque hay un nulo espacio dedicado a hablar de cómo tomar varios de estos conceptos a la vez es inconsistente[31, 48, 75, 77].

Sin embargo, en la enseñanza de la probabilidad a nivel universitario hay una insistencia en los números, en las cuentas, en los cálculos. Hay una insistencia en que el alumno pueda ejecutar cálculos estadísticos y no mucho más. Ahí mismo se puede ver que la mayoría de dichos libros no mencionan ni de pasada que existen muchas corrientes o escuelas de pensamiento sobre probabilidad y menos aún los conflictos importantes que impactan nuestro uso de la probabilidad. No es de extrañarnos, es la era que nos tocó vivir, es la ciencia que nos tocó vivir.

No somos distintos, en eso, a otros tiempos. Justo Kolmogorov vivió, y muy profundamente, la ciencia de su tiempo, la matemática de su tiempo. Una matemática que estaba profundamente enamorada del sueño Hilbertiano. Sería muy desconsiderado expornerle a usted a una explicación mía de este sueño. Baste decir que los matemáticos de principios de siglo soñaban con que todas las teorías matemáticas estuvieran basadas en axiomas. Hoy día es un poco difícil entender dicho sueño, principalmente porque se volvió en buena parte realidad. Piense usted en todos los libros de matemática de nivel universitario que empiezan escribiendo axiomas para seguir en un gigantesco rosario que en vez de padrenuestros y avemarías consta de teoremas, corolarios y, el peor de todos, problemas.

En cualquier caso Kolmogorov escribió unos axiomas que, de acuerdo a su forma de ver las cosas, capturaban la esencia matemática de la probabilidad. La capturaban dejando todos esos elementos que no consideraba esenciales para la parte matemática. Este es el primer problema. Quiero recalcar que los axiomas tenían como objetivo capturar la matemática y solo la matemática asociada a la probabilidad. Era su objetivo dejar fuera los problemas de cómo y cuándo es correcto o pertinente usar la probabilidad, cosas que trascienden sus propiedades matemáticas.

Si usted recuerda, tanto el frecuentismo, el subjetivismo, y las otras; explícitamente atacaban este problema. Justo intentaban dar esas condiciones para aplicar la probabilidad. Por ejemplo, si no hay un colectivo, no podemos hablar de probabilidades dirían los frecuentistas. Por eso los axiomas de Kolmogorov no son una teoría de la probabilidad como las que hemos estado mencionando. Intentan solo ser una regla matemática para una serie de números que llamamos probabilidades.

Es a veces difícil entender este punto, si usted está leyendo este texto muy probablemente tenga algún tipo de formación científica en la segunda mitad del siglo XX o en el siglo XXI donde preguntar por la aplicación de la probabilidad no parece tener sentido. Pero ese tipo de preguntas son válidas no solo para la probabilidad sino para cualquier teoría. Por ejemplo pensemos en la aritmética y los números naturales. Recordando a Borel, tenemos como ejemplo absurdo de meter a una oveja y a un tigre en la misma jaula y, como uno más uno son dos, esperar a la mañana siguiente encontrar dos animales. ¿Qué es lo que falló en este ejemplo? ¿Es que acaso hay algo mal con la aritmética? Claro que no, la aritmética es solo una regla matemática, una regla que se aplica a unos símbolos para, oh sorpresa, escribir más símbolos. El problema no lo tiene la aritmética sino el uso que hacemos de ella. Simplemente hay situaciones de este mundo que no siguen las reglas de la aritmética.

Análogamente, hay situaciones de este mundo que simplemente no siguen los axiomas de Kolmogorov. Eso no causa un problema en los axiomas, no es extraño, ni peculiar. Es un hecho común y corriente que pasa a cualquier conjunto de axiomas matemáticos.

Los axiomas de Kolmogorov tienen el objetivo explícito de intentan describir de la manera más abstracta posible a la probabilidad. Solo en su comportamiento matemático. Esto se puede leer directamente de él [45]:

> The theory of probability, as a mathematical discipline, can and should be developed from axioms in exactly the same way as Geometry or Algebra. This means that after we have defined the elements to be studied and their basic relations, and have stated the axioms by which these relations are to be governed, all further exposition must be based exclusively on these axioms, independent of the usual concrete meaning of these elements and their relations.

Quiero señalar que eso no significa que Kolmogorov no estuviera al tanto de los otros problemas de la probabilidad. Es decir, de esos que no tenían que ver con la descripción matemática. En esa misma página, él mismo señala la cuestión y se adscribe al pensamiento de von Mises en un comentario al pie:

> The reader who is interested in the purely mathematical development of the theory only, need not read this section, since the work following it is based only upon the axioms in §1 and makes no use of the present discussion. Here we limit ourselves to a simple explanation of how the axioms of the theory of probability arouse and disregard the deep philosophical dissertations on the concept of probability in the experimental world. In establishing the premises necessary for the applicability of the theory of probability to the world of actual events, the author has used, in large measure, the work of R. v. Mises, pp.21-27.

Es decir, él dice explícitamente que hay otros problemas en la aplicación de la probabilidad y en ellos sigue las ideas de von Mises. Esto es de crucial importancia en relación a la mecánica cuántica: aún cuando las personas estén instruidas en los axiomas de Kolmogorov, no es suficiente pues dichos axiomas no determinan el tratamiento de los problemas que presenta la aplicación de la probabilidad. Es necesario establecer un marco conceptual que permita lidiar con esos problemas, por ejemplo el marco de von Mises y su probabilidad frecuentista.

Si la axiomatización de Kolmogorov explícitamente se refiere a la teoría de von Mises en lo que se refiere a su aplicación ¿por qué no se usa esta última? Además, si no fueron realmente relevantes para las comunidades de esta era dorada de la probabilidad ¿por qué deberíamos nosotros tomarlos tan en serio en nuestra discusión sobre mecánica cuántica?

Para 1933 que se publica su monografía sobre la probabilidad, Kolmogorov ya era un matemático respetado por encabezar, junto con Khinchin, la escuela de probabilidad Rusa. Nótese también que la ciencia soviética no tenía la legitimidad que obtendría en épocas de la segunda guerra y posteriores, pero que tuvo algunos excelentes resultados en cuanto al desarrollo de los métodos de cálculo[67, 71, 3, 78].

En esta época, finales de los 20´s y principios de los 30´s, los científicos soviéticos se vieron amenazados por ser considerados parte de la burguesía. A partir de 1929 se autorizó que los profesores universitarios podían ser rescindidos de su puesto por órdenes de la administración. Esto además se suma a un ambiente en el que buena parte de las discusiones científicas de la época se consideraban de una ideología contraria al materialismo dialéctico y se buscaba eliminar. Note que antes mencionamos que la versión en ruso se publicaría tres años después de la original en alemán.

Esta cita que describe muy bien la situación a la que nos enfrenamos[16]:

> As mentioned above, Kolmogorov's publication Grundbegriffe der Wahrscheinlichkeitsrechnung is a modest monograph of 60 pages published in 1933 along with several articles devoted to the modern probability theory. The Russian translation of the text is dated 1936, and it was mostly achieved due to some political reasons at the time when an important pressure was put on Soviet scientists so that they publish their works in Russian and in USSR rather than abroad. As for the first English translation, it is dated 1950. This relatively important delay shows that the axiomatization suggested by the Russian scientist wasn't as generally accepted as we usually think it was. Several probabilists, and among the most eminent ones, such as Paul L'evy, will never use the axiomatization of Kolmogorov, which will not prevent

> them in any way having extraordinary ideas. In fact, outside the USSR, before the 50's, more or less only Cramer's treatise refers to this field. Besides, this author doesn't give any detailed explanation; he only uses Kolmogorov's axiomatization because it is the most practical one among those available at that time (in particular the theory of collectives suggested by von Mises). However, from the 50's, it will definitely be adopted by the younger generation.

Hasta antes de los 50's del siglo XX, la axiomatización de Kolmogorov era un tema marginal. Las grandes discusiones sobre probabilidad se dieron principalmente en las tres primeras décadas del siglo XX. Si bien la escuela de probabilidad soviética tuvo su lugar, fue más relacionado a temas como las nuevas formalizaciones de la ley de los grandes números, ahora conocida como la fuerte; el teorema del límite central y otros temas relacionados con la mecánica estadística.

¿Cómo es que los axiomas de Kolmogorov pasaron de ser un tema marginal para la probabilidad a estar en las primeras páginas de básicamente cualquier libro de probabilidad moderno?

Esta cuestión la dejaremos para el final de este trabajo pues su respuesta no es aislada. Que los axiomas de Kolmogorov hayan obtenido semejante estatus no es un fenómeno aislado. La mitad del siglo XX, la segunda guerra mundial, la guerra fría fueron tiempos turbulentos, muy turbulentos para la ciencia en general.

Es claro que este trabajo tiene una actitud crítica sobre la probabilidad. Por favor permítame subrayar que no se trata de criticar o de negar la consistencia matemática de los axiomas de Kolmogorov; eso no tiene sentido. Lo que se critica y se niega es la pertinencia de los axiomas para describir toda la probabilidad. Insisto toda. No solo su parte matemática. No es esto una falla de Kolmogorov, ya hemos visto que el era consiente de la situación. Es una falla mía por pecar de vanidad y hablar de Kolmogorov sin haberlo leído. También por carecer de mínimo espíritu crítico necesario como para cuestionar siquiera una palabra de todo el adoctrinamiento probabilístico que me fue impuesto en mi entrenamiento como científico.

Pequé de vanidad, pues el adoctrinamiento científico me hizo sentir que de alguna manera yo sabía algo que pocos sabían, que tenía aunque sea una autoridad mínima para hablar de probabilidades o de ciencia en general. En el fondo hay una voz, aunque sea pequeña que te recuerda lo superficial de ese conocimiento; que recuerda las muchas veces que acabé correctamente los cálculos, pero que me quedaba viendo a la página terminada, vacío por dentro bien enterado de que no entendía cuál era la conclusión ni cómo había llegado ahí.

Capítulo 5

Probabilidades al borde de un ataque de nervios

Siempre me ha sido difícil discutir sobre las probabilidades pues requiere demasiado contexto. En general, aunque ya hayamos platicado de las distintas escuelas de la probabilidad, aunque parezca interesante, no es fácil ver qué tiene que ver esto con nuestro uso de la probabilidad.

Es común seguir con la idea errónea de que nuestro uso de la probabilidad está por un lado y toda esta cuestión es una mera interpretación que está en otro lado. Así podemos pintar la caricatura donde el terreno filosófico está a nuestra izquierda, nuestro uso de la probabilidad a la derecha y los axiomas de Kolmogorov en medio como una perfecta frontera que permite con mucho éxito la división del trabajo probabilístico.

Esto no es fácil de ver a simple vista, pero tal vez un ejemplo nos ayude.

Probablemente no hay ejemplo de uso de la probabilidad más gastado que el de la moneda, seguido muy de cerca del de los dados. Imagínese, por favor, un dado. Una persona con adoctrinamiento científico se le presenta y le asigna un sexto de probabilidad a cada una de sus caras. Esto significa en nuestro lenguaje más técnico que la probabilidad del evento "el dado al ser lanzado caerá con la cara x viendo hacia arriba" es $1/6$. Hasta aquí algo cotidiano. Ahora la persona realiza una gran cantidad de lanzamientos del dado, digamos 1000. Esperamos, según nuestra idea de probabilidad, que cada cara del dado salga alrededor de 166 veces. No esperamos que sea exacto, pero sí que la desviación de 166 sea pequeña.

Continuemos con el ejemplo, supongamos que después de los lanzamientos no obtenemos lo esperado. Digamos que el número 3 sale 500 veces. Así que pensamos que la cara de ese dado está cargada. Después de mucho tirar el dado concluimos que el 3 tiene 1/2 de probabilidad de salir, mientras que el resto de los números tienen 1/10 de probabilidad de salir.

Aquí viene la parte interesante: ¿Cuál es la probabilidad de que salga 3? antes de que se realice el primer lanzamiento? Es decir, cuando todavía no se tiene la evidencia de que el 3 sale más ¿cuál es la probabilidad?

Desde la aproximación clásica/logicista la probabilidad de que salga 3 realmente es 1/6 según el principio de indiferencia. Conforme el experimento se va realizando, nueva evidencia se obtiene. Esto va cambiando el valor de la probabilidad y se llega al final al valor de 1/2, pero quiero resaltar que el valor antes de los experimentos sí es 1/6.

Desde la aproximación frecuentista la respuesta es muy diferente. Esta es una probabilidad objetiva y no depende de la evidencia disponible en el momento. Antes de tirar el dado 1000 veces la probabilidad ya era 1/2, cuando hicimos el experimento solo "revelamos" el valor, lo "descubrimos". El valor no cambió durante el experimento, siempre fue 1/2.

Querer tomar este par de aproximaciones al mismo tiempo es algo común en el uso actual de la probabilidad. Se suele asociar probabilidades usando la probabilidad clásica, pero se habla de ellas como si fueran propiedades objetivas, es decir se hablan de ellas como si se hablara de masas, velocidades o distancias.

No creo que le sea difícil pensar en ejemplos donde hacemos esto cotidianamente. Pero como resulta que sí es difícil pongamos algunos.

Pasemos a la otra propiedad importante para nuestra cuestión.

Seguro en algún momento ha escuchado que al tirar un par de dados, y según la costumbre sumar sus resultados, el número más común en salir es el 7. Olvidémonos un momento si es una probabilidad epistémica u objetiva. Solamente digamos que el 7 sale más. Ahora yo me encuentro en un juego de azar, apuesto dinero, una buena cantidad y como me indica mi adoctrinamiento científico le apuesto a que sale 7. Oh sorpresa, sale 10 y pierdo todo mi dinero.

Como consuelo, todas las personas que me rodean, que también les es familiar la probabilidad dicen que tomé la "mejor" decisión. Mis maestros de probabilidad dicen lo mismo. Los libros de probabilidad dicen lo mismo. Pero ¿cómo lo saben? ¿Cómo podemos saber que es la mejor? ¿en qué sentido es la mejor?

Cuando intentamos responder estas preguntas nos topamos con dificultades. Uno puede responder que es la mejor porque es la que la probabilidad dice. Esta respuesta no es la que buscamos ¿no? pareciera una falacia de autoridad. Otra respuesta más común es que es la mejor apuesta porque si realizáramos el juego muchas veces, en la mayor parte saldrá 7. Luego hay que apostarle al que más sale ¿no?

Pero hay un gran problema con esta respuesta, tiene una condición, "si realizáramos el juego muchas veces" este es el problema. La apuesta sólo la voy a hacer una vez. Solo una. No tengo el dinero para apostar 1000 veces, solo tengo para una. ¿Por qué no apostarle al 8 que es lo que me recomendó el tarot, los pozos de té, la astrología o cualquiera de esas prácticas que la ciencia condena como charlatanería?

Si sólo voy a hacer la apuesta una vez, si solo voy a ir al programa de concursos una sola vez, con qué derecho la ciencia y la probabilidad podrían decir que su apuesta es "la mejor", que está por encima de esas "terribles supersticiones" como el espiritismo o simplemente que desde el fondo de mis entrañas sentí que el número que saldría es el 10.

Sí hay una respuesta en nuestras comunidades científicas para eso, pero en general es que "es la más probable", pero esta respuesta tiene justo el problema de la primera respuesta: tenemos una especie de argumento de autoridad: es la mejor apuesta porque la probabilidad lo dice.

Este tipo de situaciones son comunes en el uso de la probabilidad, tomamos algo que se aplica a un grupo de eventos y lo extrapolamos al caso individual. Podemos argumentar que la mejor estrategia en el juego discutido es escoger el 7, siempre que se juegue muchas veces. Para hacer una línea argumental que justifique esta estrategia para el caso donde solo juego una vez debo extrapolar del caso grupal y esto no está justificado.

Estas situaciones se presentan en cosas cotidianas, realmente cotidianas. Uno de los mejores ejemplos es el del señor fumador a quien su médico le recomienda dejar de fumar. 9 de cada 10 personas que fuman como usted fallecen. Le dijo. La probabilidad de que usted fallezca de enfermedades relacionadas con el tabaco es muy alta, de 9/10. Le dijo. ¿Es este un argumento para dejar de fumar? Al fin y al cabo el es sólo una persona, con solo una vida en juego. Aunque dichas estadísticas fueran ciertas ¿cambia en algo para la decisión *individual* de dejar de fumar?

Este es un tema de gran importancia para la probabilidad. De nuevo si se le pregunta a las distintas escuelas tendremos distintas respuestas. La aproximación clásica/lógica y la subjetiva dirá que sí aplica al caso de eventos individuales. La frecuentista dirá un rotundo no, no es posible aplicar la probabilidad a eventos individuales. La probabilidad solo aplica a grupos de eventos, eso que llaman el colectivo. Las propensiones, como ya se mencionó, justo intentarán dar una probabilidad objetiva, pero que se aplique al caso individual. Sería difícil decir que lo lograron, el porqué es muy interesante.

5.1. El problema de la clase de referencia

Decidir si la probabilidad es objetiva o epistémica y aplicarla a un grupo de eventos o a un evento individual es probablemente el problema más importante del pensamiento probabilístico moderno. Sus implicaciones llegarán a todo el uso de la probabilidad en la actualidad y por extensión será fundamental en el entendimiento de la mecánica cuántica.

En esta sección hemos insistido dos dicotomías que tiene el pensamiento probabilístico. Primero la probabilidad puede ser objetiva o epistémica. Objetiva en el sentido que es una cantidad objetiva, de la misma manera que la masa o la distancia. Se asocia a objetos o a situaciones y pensamos que es una cantidad independiente de qué pensemos, de qué información está disponible para nosotros o de cuales son nuestras creencias. Tenemos dos ejemplos de dichas teorías el frecuentismo y las propensiones. Epistémicas son esas teorías de la probabilidad que sí dependen del pensamiento de las personas, donde la probabilidad puede cambiar solo porque obtenemos nueva información o simplemente porque nuestras creencias cambian. Tenemos también dos ejemplos: la teoría logicista y la subjetivista.

La segunda dicotomía es la de la aplicación de la probabilidad el caso singular o individual. Hasta ahora solo lo hemos discutido para el frecuentismo. En dicha teoría es de fundamental importancia que la probabilidad solo se aplique al caso de colectivos, aplicarla al caso individual no tiene sentido en dicha teoría. ¿Por qué insisto en esto? ¿Por qué es tan importante que no se use?

Supongamos que dentro de el espíritu del frecuentismo intentáramos dar un sentido a la probabilidad de un solo evento. Recordemos que el frecuentismo intenta representar un “hecho empírico”, a saber que en un conjunto grande de eventos encontramos estabilidades en las frecuencias relativas. El problema es que solo tenemos un evento y no hay tales frecuencias. Luego buscamos un colectivo donde poner nuestro evento para asociarle una probabilidad. Pero hay una infinidad de ellos, hay una infinidad en los que este evento se puede poner. Luego tendremos que decidir, ciertas comunidades escogerán alguno y otras otro. Ahí se perderá la objetividad de la probabilidad.

No tenemos una forma de escoger un colectivo que represente un solo evento. Ese evento puede compartir características con una infinidad de colectivos. Vamos a repetir lo mismo pero hablando de monedas. Lanzamos una moneda, ¿cuál es la probabilidad? Solo hubo un lanzamiento. Digamos sale Águila. Ese evento se puede meter fácilmente en un colectivo de lanzamientos que sean 50 % Águila y 50 % Sol. También se puede meter en

$40-60$, $30-70$ y básicamente cualquier combinación. Incluso $100-0$ para sorpresa de quienes no aplican fielmente la definición de von Mises. Luego se puede meter un lanzamiento de águila a cualquier colectivo, que corresponda a cualquier valor de probabilidad. Todos los colectivos son igual de válidos para meter a este lanzamiento.

Esto representa el problema fundamental para poder asociar una probabilidad frecuentista a un fenómeno individual.

Voy a continuar esta discusión forzándola un poco para el público dedicado a la física. Normalmente, quienes tenemos un entrenamiento en la física mina nuestra forma de pensar este mundo y al encontrarnos con este problema intentamos darle la vuelta al problema haciendo uso de nuestra formación.

Por ejemplo, en el caso del lanzamiento de la moneda, contestamos por acto reflejo que el colectivo de lanzamientos que le corresponden al lanzamiento de esa moneda, es el colectivo formado al lanzar *esa* misma moneda un gran número de veces. Luego es fácil que abandonemos el punto repitiendo, también por reflejo, la discusión del determinismo, pasando forzosamente por el rezo de siempre: las condiciones iniciales, etc.

En general, la idea se centra en alguna idea vaga de las propensiones, es decir, asociar la probabilidad a las condiciones en las que un evento pasa. En el caso de la moneda queremos que la moneda sea lanzada de nuevo con ciertas condiciones. Tratamos de controlar que sea la misma moneda aunque aceptamos que sea lanzada de una manera ligeramente diferente, donde ligeramente es un criterio totalmente subjetivo.

Vamos a ahondar mas en la cuestión. Recordemos que lo haremos en el marco de las propensiones y será en las propensiones a largo plazo. Lo importante por ahora es que las propensiones son una propiedad de las condiciones en las que un evento se lleva a cabo. La estrategia para librar el problema de la clase de referencia sería entonces que al evento individual se le asocie la probabilidad de las condiciones en las que el evento se llevó a cabo. Esta era la idea original de Popper al plantear las propensiones.

De nuevo tenemos el problema de clase de referencia. En las propensiones podemos hacer la pregunta "¿Cuál es el conjunto de condiciones que deben caracterizar a nuestro evento individual?". Si tomamos el ejemplo de la moneda podemos pensar que si se lanzó con la mano derecha o la izquierda no debería en principio ser importante para la caracterización del lanzamiento, pero bien puede ser el caso que al lanzarla muchas veces con la derecha y muchas veces con la izquierda, nos encontremos con diferentes resultados. Bien puede ser que lanzarla en martes da diferentes resultados que lanzarla en sábado y así con muchas otras propiedades. La respuesta natural es decir que si nuestro lanzamiento fue con la mano izquierda un sábado, tenemos que tomar en cuenta esas condiciones. Pero este camino no tendrá fin. Siempre podremos agregar más condiciones y más condiciones hasta que nos quedemos con tan pocos eventos como para medir una frecuencia relativa. Es decir, en el ejemplo de la moneda, tenemos que agregar condiciones como un lanzamiento con la mano izquierda, en sábado, a 29 grados centígrados de temperatura ambiente, a las 9 de la mañana, tomando en cuenta el color de los calcetines que usamos, el estado de la bolsa, etc. Esa lista seguirá hasta que sólo nos quedemos con un lanzamiento y no habrá frecuencias relativas pues solo será un lanzamiento.

Así es como Gillies propuso su clasificación de las teorías de propensiones[36]. Las clasifica dependiendo de cómo enfrentan el problema de la clase de referencia. Si aceptan que las propensiones son propiedades de las condiciones de un evento y que si se tienen muchos eventos en esas condiciones obtendremos unas frecuencias relativas estables. Esta aproximación renunciará a que se puedan asociar probabilidades objetivas a los evento individuales, al igual que en el frecuentismo se puede asociar un evento individual a un gran número de posibles conjuntos de condiciones y elegir unas condiciones es subjetivo.

Por otro lado, se puede proponer que cada evento individual se debe caracterizar con las condiciones más "completas" más "universales" posibles. Aceptar que esas condiciones tienen una propensión asociada, que naturalmente es inmedible pues esas condiciones no se pueden repetir, y pensar que las combinaciones de esas propensiones son las que terminan dando los valores de que se miden con frecuencias relativas. Esta segunda opción no renuncia a la objetividad de las probabilidades en los eventos individuales, pero lo hace a un costo muy grande: las probabilidades de un fenómeno individual dependerán de las condiciones universales en las que un evento se lleva a cabo. Probablemente si usted está relacionado con la física esto le traerá algún recuerdo del experimento EPR o las interminables discusiones alrededor de Bell. Tocaremos ese punto más adelante.

5.2. La probabilidad que uso todos los días no existe

"Nuestra vida se sostiene sobre discursos oficiales históricamente falsos" diría Carlos Fuentes a finales de los 70 en su defensa del regimen de Echeverría. Toda proporción guardada, el discurso científico moderno vive sobre una idea falsa de la probabilidad.

Si soy sincero y repaso las ideas de la probabilidad en las que fui educado encuentro una probabilidad que no existe. Cotidianamente hablo de la probabilidad como si fuera una cantidad objetiva, al mismo tiempo le asocio esas mismas probabilidades a fenómenos individuales. Pero ignoro las condiciones universales, tomo solo condiciones parciales para asociar probabilidades. Esto es contradictorio, el problema de la clase de referencia me prueba todos los días que lo que digo no tiene coherencia.

Cuántos más al igual que yo, y de maneras mucho más masivas y brutales, hacen lo mismo. Me gustaría decir que son pocos, pero no puedo. Realmente, en mi vida, he encontrado expertos en las ciencias que todos los días usan el mismo discurso que yo. Igual de ignorante e igual de incoherente y así ejercen la ciencia.

No hay una probabilidad que cumpla todas esas cosas que pedimos. No tiene sentido hablar de una probabilidad objetiva, que se aplique al caso individual y que no dependa de unas condiciones "universales". Estos tres requerimientos no se pueden cumplir. Quiero que quede muy claro que no se pueden cumplir.

Le invito a pensar en las ingentes cantidades de novedosos artículos científicos que todos los días son publicados. Independientemente de la exclusividad de la revista donde salen, cuántos de ellos hacen uso (o abuso mejor dicho) de una idea de la probabilidad con justo las propiedades incompatibles que acabamos de mencionar.

Le invito a pensar un momento en todos los usos que usted hace de la probabilidad y que *coram Deo* se diga cuales de esos usos tienen sentido. Ejercicio interesante aunque encontremos cosas desagradables de nuestra idea de la ciencia, la racionalidad o incluso de nuestro pensar y conducir por este mundo.

Yo hice este ejercicio y le confieso encontré más palabras que verdades.

Capítulo 6

Intermedio

Antes de continuar y llevar nuestra discusión sobre la probabilidad a la física, me gustaría hacer una pausa para decir unas palabras sobre el estudio de la mecánica cuántica y los teoremas no-go en relación a la probabilidad. Quizás lo más sorprendente de atacar estos temas no son los temas en sí, sino la literatura que no parece tener ningún orden ni sentido común en su desarrollo.

Quizás al grueso de los científicos no nos parece natural estudiar la probabilidad al estudiar mecánica cuántica. ¿Por qué? Hay mil palabras clave que se repiten hasta el cansancio en la literatura actual, que localidad, que contextualidad, cuántico, aleatorio, indeterminismo, valor indefinido, quantumness, value-indefiniteness, solo Dios sabe la infinita variedad en la que estos términos se usan y, sobre todo, se venden. Los talegos de experimentos que francamente agotan posibilidades risibles de "aleatoriedad" a la hora de elegir las bases en las que se harán las mediciones en los experimentos de Bell.

El sinsentido parece gobernar cuando vemos experimentos donde se usan "decisiones" humanas para elegir las bases, según, dicen, intentando darle un revés al super-determinismo[6], pero semejantes intentos parecen la carrera armamentista de la necedad y la superstición científica. Que no veamos una análisis profundo del concepto de probabilidad, necesario para plantear todas estas ideas, antes de lanzarse al abismo de la conjetura que es el super-determinismo o a la proeza quijotesca que representa un experimento donde se eligen las bases con humanos, me parece un sinsentido. Quizás sea yo el bruto, pero es que después de escribir incontables veces la letra P en nuestros "calculitos" ¿nunca se nos ha ocurrido preguntar qué significa?

De hecho todos los científicos, incluso fuera de la mecánica cuántica saben de la probabilidad, es más, toda la gente que pertenece a esto que el mundo globalizado considera mundo, ha oído de la probabilidad. Tampoco es difícil de creer que el grueso de aquellos quienes nos reconocemos como científicos venimos de ese mundo, al fin y al cabo el mero hecho de pasar por el adoctrinamiento científico en alguna institución universitaria ya parece suficiente.

La familiaridad de las personas con este concepto realmente es un asunto digno de estudio. Sin ahondar en el tema, que ya es bastante viejo dicho sea de paso, es natural pensar que alguien con mucha familiaridad con la probabilidad no tendrá ningún reparo al usarla en sus infinitas variedades para intentar "ponerle números al mundo". Es bien conocida la crítica de Heidegger a los conceptos científicos en el lenguaje cotidiano. Dicha introducción los termina por hacer "dogmáticos" de alguna manera. Hoy día es algo que se puede constatar con mucha facilidad, y según mi visión con algo de tristeza.

Solo para poner un ejemplo, podemos consultar los libros de electricidad y magnetismo de finales del siglo XIX y principios del XX y constatar que se menciona la hipótesis de la carga eléctrica. Es decir, se afirma claramente que se debe suponer la existencia de la carga eléctrica a lo largo del libro y se dan razones que "motivan" dicha suposición. Los libros de la materia en la segunda mitad del siglo XX no solo aceptan dicha hipótesis sino que se van de boca y sencillamente no la mencionan. Dicha noción no causa inconformidad en el grueso de los estudiantes al fin y al cabo la mayoría ha ya escuchado sobre la carga eléctrica e incluso sobre sus propiedades, que hay dos tipos de carga, etc. Así que dichas propiedades parecen estar más del lado del dogma científico. Pregunte usted a sus colegas científicos cuál es la evidencia más contundente de que existen dos tipos de carga y se encontrará que en general no le podrán mencionar ni una.

La probabilidad entonces adolece también de ese dogma. Expertos en estadística, mecánica cuántica y otras disciplinas que se basan fuertemente en la probabilidad, suelen no estar al tanto de problemas básicos que imponen límites a la probabilidad. La mayoría no ha escuchado siquiera sobre el problema de la clase de referencia que parece el problema más relevante del pensamiento probabilístico moderno. De esta manera no somos más que meros usuarios de un discurso inconsistente e ignoramos su inconsistencia. De hecho no solo la ignoramos sino que, ejerciendo un cientificismo brutal, creemos ciegamente en que dicho discurso es consistente, impoluto y eterno.

Qué ironía de la vida, durante mi educación como físico recuerdo los mil y un comentarios que tanto profesores como alumnado realizábamos denostando a la ingeniería como disciplina. Se señalaba siempre que de alguna manera la ingeniería hacía un uso ciego de las "ecuaciones" o las "leyes físicas" y que el estudio que realizábamos nosotros se salvaba de eso. Pareciera que dicha situación trae irremediablemente reminiscencias de sobre-compensación psicológica. Como si insistiéramos tanto en esto porque probablemente en el fondo todos sabemos que nuestro conocimiento de las leyes y las ecuaciones era igual de superficial, frívolo e instrumental que el de la ingeniería; que al escribir ecuaciones nosotros también ignorábamos el significado de casi todos los conceptos y solo estábamos condenados al igual que una computadora a manipular infinitamente símbolos que jamás entendíamos.

Volviendo al tema principal, dicha fe ciega en la probabilidad será la clave para nuestro análisis de la situación en mecánica cuántica. ¿Cómo podríamos darle sentido a la ecuación de Born cuando no sabemos qué significa la P que usamos en dicha ecuación? Ahora, permítame hacer un paréntesis. No tiene nada de malo ni condenable plantear ecuaciones que no tenemos ni idea que significan. De hecho es algo que se hace constantemente en la ciencia. No solo en esa ciencia ya gastada de hoy, sino incluso en sus mejores tiempos. Póngase por ejemplo la introducción sin fundamento de la μ que Gibbs usó y que después se conocería como el potencial químico, o el *spannung* que Ohm utilizó en sus investigaciones sobre la corriente y así un sin fin de ejemplos. No tiene nada de malo ni de condenable.

Sin embargo, no podemos pedir que una gran consistencia a ese discurso. Estos planteamientos son muy buenos porque permiten, dicho abiertamente, atinarle a los datos de maneras y con precisiones increíbles en ese contexto. Milagros científicos, dirían aquellos que ignoran la historia de su disciplina. Pero esas proezas no ayudan a dar sentido y coherencia a una teoría.

Toda proporción guardada vemos lo mismo en el uso de la probabilidad. Qué bueno es ese uso inconsistente de la probabilidad para atinarle a los datos, pero qué inconsistente sale a la hora de proponer explicaciones. La mecánica cuántica es igual.

Quién que esté cerca de estos temas no he escuchado decir a los cuatro vientos esa bandera del cientificismo moderno que versa "la mecánica cuántica es la descripción mas precisa que se ha tenido. . . " y puede usted completar como desee la frase. Bandera que viene acompañado de su revés, que ese sí no se grita a los cuatro vientos, como podemos ver en las aseveraciones de Sakurai en su famoso libro de mecánica cuántica donde dice que "When the physical system in question has no classical analogues, we can only guess the structure of the Hamiltonian operator. We try various forms until we get the Hamiltonian that leads to results agreeing with empirical observation"[69, pág.85]. Y es que este es punto curioso ¿cómo no va a ser tan exacta si siempre nos andamos inventando hamiltonianos con tal de atinarle a los datos? ¿cómo no le vamos a atinar con tanta precisión si, desde las eras del espín[57, 74], hemos inventado propiedades físicas risibles con tal de atinarle a los datos?

Por más goloso que parezca, no entremos al problema de la demarcación y dejemos de lado esas aseveraciones. Al final, el problema es que usamos la probabilidad como siempre lo hemos hecho, como hemos sido adoctrinados a usarla, y dogmática o no, ese uso es incoherente, es inconsistente. Le atina muy bien a los datos, pero no podemos esperar que nos dé un discurso bien formado. Si queremos consistencia pues será lo más natural que partamos de un discurso consistente de la probabilidad. Ese el camino que le propongo a usted para los siguientes capítulos.

Para finalizar repetiré el que es el mensaje más importante de este trabajo: nuestro uso de la probabilidad es incoherente, inconsistente, dogmático e ignorante. Como científicos nos vanagloriamos de la lógica, la razón, el pensamiento crítico; bueno pues en el terreno de la probabilidad o es mentira o en el mejor caso hipocresía. ¡Qué bien a los datos! nadie lo discute, que nuestro discurso es incoherente ¿Cómo negarlo?

Hay quienes piensan que eso aplica a toda la ciencia, yo soy uno de ellos, pero no es aquí el lugar para discutirlo.

Parte II

Vanidad cuántica

Capítulo 7

Los nuevos colapsos del emperador

Bajo la luz que la visión crítica e historiográfica de la probabilidad nos ofrece, problemas tan complejos y endiosados como el del colapso de la función de onda empiezan a perder un poco su brillo y majestuosidad.

Pensar que la función de onda presenta algún tipo de proceso que lo lleva de una "superposición" lo que sea que eso significa, a un "valor definido", también lo que sea que eso signifique, es una idea presente desde la mecánica de Schrödinger. Para la segunda mitad de la década de los 1920 y la de 1930, encontramos una buena cantidad de trabajos con diferentes visiones del problema.

Comenzaremos con una pregunta que me parece la más indicada para plantear el problema que tratamos en este trabajo. La pregunta es la siguiente:

Hay una infinidad de trabajos lidiando con el colapso, pareciera que como comunidad nos molesta. Parece que nos molesta que el valor del estado cuántico pase cambie de manera "instantánea", que se pase de una superposición a un autoestado así sin más. Sin embargo, yo no nos veo como científicos, en particular como físicos, saltar de sorpresa cuando la probabilidad parece hacer lo mismo ¿cómo es que si el estado cuántico cambia de valor todo mundo se lleva las manos a la cabeza, pero todos los días tomamos probabilidades con ese mismo comportamiento y ni siquiera levantamos una ceja? ¿Acaso no debería haber un problema del colapso de la probabilidad?

Antes de lanzar una moneda, es costumbre decir que la probabilidad es un medio, ¿cuál es la probabilidad de que salga Sol en el volado una vez que ya salió Sol? ¿es uno? Y si es uno, por qué no reclamamos que cambió de

1/2 a 1 instantáneamente y que la probabilidad de Águila pasó de 1/2 a 0 instantáneamente.

Estos elementos son suficientes para plantear nuestra cuestión. Qué derecho tenemos nosotros de decir que, en un volado, la probabilidad de que caiga Sol es un medio. Respuesta ninguno. Por más que sea un dogma enraizado en el más profundo corazón de la ciencia y los científicos, no tenemos grandes bases para poder asegurar dicha asignación.

Según las escuelas de la probabilidad, esta asignación no está justificada porque faltarían mediciones, o simplemente no tiene sentido porque el solo hay un lanzamiento y la probabilidad se asigna a colectivos y no a eventos individuales; o porque el lanzamiento es uno y la probabilidad se asigna a condiciones las cuales no están bien especificadas si solo se hace referencia a un lanzamiento. Estas son las razones por las que en el subjetivismo, el frecuentismo y las propensiones de largo plazo dirán simplemente no tenemos derecho a hablar de dicha asignación.

Las únicas corrientes de pensamiento sobre la probabilidad que aceptan dicha asignación son la clásica y la logicista. Que son las más emparentadas entre la variedad de escuelas que existe. Dicha asignación basada o en el principio de indiferencia o el de razón insuficiente, respectivamente tiene sus problemas como ya se ha discutido. Solo en esas se puede aceptar una asignación de este tipo.

Eso es lo que se puede decir sobre asignar un medio a el caso “cayó Sol” antes de lanzar la moneda.

Después de lanzar la moneda la historia es un poco diferente. De hecho, si bien nuestro corazón de ciencia grita “un medio” antes de lanzar la moneda, las respuestas suelen ser más variadas cuando se pregunta la probabilidad después de haberla lanzado. Esto revela una falta de adoctrinamiento uniforme en la comunidad científica. ¡Claro!, en los libros de probabilidad hay pocos problemas donde se pregunte la probabilidad de algo una vez que ya pasó. Sin embargo, la probabilidad de que salga Sol una vez que ya salió Sol en muchos casos se toma como 1.

De igual manera vamos a preguntar con qué derecho decimos que la probabilidad de que salga Sol una vez que ya salió sol, es 1. La respuesta, para bien o para mal, es de hecho igual de vacía. Dependerá de la corriente de pensamiento que adoptemos. Por ejemplo el subjetivismo sí nos permitirá decir que la probabilidad es uno, pues en general se puede decir que las personas creerán que salió uno. Aunque siempre es posible que haya personas que no lo crean aunque lo vean y simplemente sus creencias no entraran dentro de nuestros estándares de racionalidad(dígase axiomas de Kolmogorov). Desgraciadamente para las propensiones y el frecuentismo la historia no habrá cambiado mucho. Sigue sin haber un colectivo ni una especificación de las condiciones, así que de ese lado seguimos igual. Curiosamente de lado del los clásicos/lógicos sí hubo un cambio pues dada la evidencia de que salió Sol se infiere que salió Sol, tautología de manual.

Hasta aquí solo hemos dicho que: en el subjetivismo y en la teoría clásica/logicista sí se puede, aunque sea a veces, argumentar que esas asignaciones de probabilidad a las que estamos acostumbrados son correctas. Por otro lado, hemos dicho que esas mismas asignaciones en las propensiones o en el frecuentismo no son correctas. Mejor dicho no tienen sentido pues ni siquiera hemos trabajado con el objeto más importante de cada teoría, (para el frecuentismo el colectivo, para las propensiones las condiciones)

Sin embargo es posible que las personas quieran decir que si bien no hablamos de colectivos y condiciones, pudimos haberlo hecho. Así que vamos a hacer una pequeña digresión para pensar como podemos asignar probabilidades en el frecuentismo y las propensiones.

Para asignar probabilidades en el frecuentismo necesitamos un colectivo de monedas que se lancen. Imaginemos que las lanzamos todas e increíblemente caen exactamente la mitad en Águila y la mitad en Sol. Si tomamos una moneda más y preguntamos cuál es la probabilidad de que caiga Sol antes de lanzarla, nos sentiremos con toda la autoridad de decir que es un medio. Ese es el método que siguen los frecuentistas. Hasta ahí todo bien.

Ahora la lanzamos y cae Sol. Así que preguntamos cuál es la probabilidad de que salga Sol después de que ya cayó Sol. Tristemente es aquí donde nuestra respuesta estándar será incorrecta. La probabilidad sigue siendo $1/2$, o en tal vez $1/2 + \epsilon$ si se quiere considerar este último lanzamiento.

La probabilidad en el frecuentismo no se puede aplicar a ese evento solo y se tiene que aplicar a todo el colectivo donde ya había caído la mitad águila y la mitad Sol, por eso la probabilidad sigue siendo $1/2$. Así que la probabilidad, no cambió, o en todo caso será un cambio muy pequeño e idealmente cero pues el colectivo es infinito en principio.

Las propensiones tienen una historia similar y mucho más sencilla. Si especificamos unas condiciones, digamos las escribo, y logro concluir que la propensión de que caiga Sol es un medio, diré que el valor es un medio. Si la lanzo y cae Sol no podré decir que la probabilidad es 1, sigue siendo $1/2$ pues la especificación que hice de las condiciones tampoco cambió. Es la misma. El valor de la probabilidad, cualquiera que sea, es el mismo.

En ambos casos la probabilidad no cambia(o cambia de manera despreciable) pues la probabilidad ya no está ligada a algún evento individual sino a algo más estable, ya sea el colectivo o las especificación de las condiciones donde se llevan a cabo los eventos. En cualquiera de los dos casos la probabilidad no cambia.

Así que para marcar un primer paso en nuestra cuestión original, diremos que en la probabilidad frecuentista y en las propensiones no tenemos un análogo del colapso pues la probabilidad no "salta de valor" como sí decimos que lo hace el estado cuántico. Es solo en el caso de la probabilidad subjetiva o la escuela clásica/logicista que sí se puede argumentar que hay un cambio repentino en el valor de la probabilidad.

Ahora sí pasemos al quid de la cuestión. Resulta que en esas donde sí podemos hablar del colapso de la probabilidad, la probabilidad no es un elemento objetivo. Seguro usted lo habrá notado. La escuela subjetivista, reafirmamos, dice que las probabilidades son una medida de la intensidad con la que una persona cree en un enunciado. En este caso que tan fuerte cree en que saldrá Sol al lanzar la moneda. No parece nada descabellado que antes de lanzar la moneda la intensidad con la que alguien creen en que va a caer Sol sea del 50 % y que una vez que ya cayó Sol esa intensidad vaya inmediatamente al 100 %. Es más, aunque nos dijeran que este cambio es instantáneo, que pasa como por arte de magia de un valor a otro seguiría sin tener nada de raro. Si me entero 10 años después de que la moneda realmente se lanzó y hasta ese instante el valor pasa de 50 % a 100 % sigue sin tener nada de raro.

Seguro usted ya se habrá dado cuenta que todos esos símiles son preguntas que se realizan cuando la misma situación se analiza dentro del marco de la mecánica cuántica y se suelen mostrar(o mejor dicho vender) como características muy raras y de alguna manera atractivas(entre más complicado más interesante) que tienen este tipo de fenómenos. En el caso de la probabilidad subjetivista nada de esto es algo que va bastante de la mano con el sentido común.

El caso de la logicista/clásica es similar. La probabilidad representa el grado en que un conjunto de enunciados, considerados como evidencias o hipótesis, implica un enunciado. Al agregar el enunciado "cayó Sol" cambia dicho conjunto y el valor con el que el enunciado se ve implicado cambia. De nuevo nada de diferente. De hecho estrictamente el valor de probabilidad no cambió sino que nosotros cambiamos el conjunto de hipótesis. Así que ¡cómo no va a cambiar el valor si la pregunta está cambiando!

En estos casos la probabilidad cambia pero no nos lleva, como en el caso del colapso, a que se anuncie como un misterio del universo. La clave para entenderlo es simple: en este caso la probabilidad depende de las creencias de una persona o de la evidencia disponible. Cualquier cambio en la probabilidad se explica con elementos de la vida diaria: una persona cambió de creencias(que en todo caso este sí es un gran misterio del universo) o la evidencia cambió que efectivamente es algo que esperas cuando nos enteramos de algo ¿No?

Desgraciadamente este es el callejón sin salida a que lleva la cuestión que planteamos el principio: No hay colapso en la probabilidad. En las probabilidades objetivas, esas donde el valor de la probabilidad depende de algo objetivo, independiente de la persona, el valor simplemente no cambia. Caiga Sol o no, el valor no da ningún salto. No hay ningún colapso. Para las probabilidades epistémicas, esas que sí dependen de las creencias o de la evidencia que una persona tiene, ahí sí cambia el valor de la probabilidad, pero ese cambio se explica simplemente porque lo que cambió es o una creencia o una evidencia. La probabilidad no es ninguna "variable física" que de repente haya pasado de valer h a $2h$ ni nada por el estilo. La probabilidad no es como la longitud o el momento angular. La probabilidad no habla de ese mito llamado "mundo objetivo", habla de creencias y de enunciados. No hay nada más que decir.

Somos nosotros los que hemos creído, de manera consciente o no, que hubo un cambio. Somos también nosotros los que decidimos creer que ese cambio era algo del mundo "físico", "real" "objetivo" o cualquier otro término intrincado.

Cómo entonces es que el estado cuántico tiene ese "colapso". No solo eso, ¿cómo es posible que se sostenga la regla de Born si dicha ecuación tiene de un lado al estado cuántico y del otro una probabilidad? Si una tiene un salto repentino la otra debería también tenerlo ¿no?

Ahora sí hablemos del colapso de la función de onda.

Capítulo 8

La regla de Born

Hasta ahora no hemos hablado mucho sobre mecánica cuántica. Hemos caminado por los fundamentos de la probabilidad, las contradicciones de nuestra ciencia, pero no hemos tocado el tema principal.

Primero debemos aclarar que a lo que hoy nos referimos como mecánica cuántica no es, ni de cerca, lo que el término refería en la edad de oro de la ciencia. No hace falta mas que leer superficialmente el libro de Dirac o el de Von Neumann, considerados las primeras formalizaciones de la mecánica cuántica, para darse cuenta de la mucha distancia que hay entre esas ideas y las que se pueden encontrar en los libros de texto actuales, dígase Griffiths a nivel básico, dígase Cohen a nivel mas complejo. Así que hace falta especificar al menos un poco sobre nuestra visión de al mecánica cuántica.

Nos limitaremos a hablar solo de una pequeña parte de la mecánica cuántica: la regla de Born. Tomamos esta porque es la única que se relaciona explícitamente con probabilidades. Dicha ley tiene varias formas de expresarse dependiendo del discurso con el que se presenta. Aquí utilizaremos:

$$|\langle \xi \,|\, \psi \rangle|^2 = P(\Xi = \xi) \tag{8.1}$$

Como es costumbre, se acompaña a la ecuación anterior con una jaculatoria diciendo que la probabilidad de lado derecho es la probabilidad de que al medirle la magnitud Ξ al sistema en cuestión obtengamos el valor ξ. De lado izquierdo tenemos $|\psi\rangle$ que es el estado del sistema y $|\xi\rangle$ que representa el eigenestado que corresponde al eigenvalor ξ.

De lado de la probabilidad, como seguramente usted sabe, no se especifica qué tipo de probabilidad se usa en esta ecuación. De lado del estado cuántico, difícilmente se especifica qué se quiere decir con el estado cuántico.

Para describir el estado cuántico no es difícil encontrar frases como "toda la información del sistema" y otras. Pero difícilmente una descripción de este tipo aclara la situación. Pensando en un electrón ¿el sistema es realmente solo un electrón o son muchos electrones que consideramos iguales? ¿A que se refieren con información? ¿Desde cuándo un sistema guarda información? ¿Cómo sabemos cuando tenemos "toda" la información?

En la mecánica newtoniana no había semejante frase. En dicha mecánica se hacía una descripción de algunas propiedades de los objetos, normalmente de su posición y velocidad y aceleración y masa y muchas otras. Además, al menos originalmente, no se dijo que eso era toda la información, ni parece necesario para desarrollar la teoría.

No creo que sea necesario seguir insistiendo que sobre el estado cuántico siempre hay muchas palabras y poca claridad. Así que le pido nos fijemos solo en una cuestión sobre el estado cuántico: si el estado cuántico es una propiedad objetiva o epistémica.

Los términos anteriores los usamos en el mismo sentido que los usamos para describir los diferentes tipos de probabilidad. Es decir, el estado es objetivo si no depende del conocimiento que tenemos, solo depende del sistema mismo. En cambio el estado es epistémico si depende del conocimiento, evidencias o creencias. Como ya hemos dicho antes la longitud de una mesa es una propiedad objetiva, la mesa mide lo mismo no importa el conocimiento que tengamos de la mesa, si conocemos o no su longitud, su masa, si es de madera o de metal, se acepta que la mesa mide lo mismo. Si estamos dispuestos a subirnos en la mesa es una propiedad epistémica, dependerá de lo que nosotros conocemos de la mesa. Si conocemos su composición, si nos hemos subido antes, qué tan alta nos parece, si hemos visto que puede sostener mucho peso, si representa un peligro una caída, etc.

8.1. Una pequeña aclaración sobre el colapso

Con colapso, la comunidad normalmente se refiere al cambio que se le asocia al estado cuántico al realizar una medición al sistema. Esto es problemático cuando no tenemos en cuenta las propiedades que discutimos anteriormente para la probabilidad. Es decir, si se aplica a individuos o a grupos, si es epistémica u objetiva. Vayamos paso a paso.

Primero quiero especificar que el proceso que llamamos colapso toma lugar cuando forzamos al estado cuántico a ser objetivo y al mismo tiempo aplicarse a individuos. Digamos que tenemos un sistema al que le asociamos un estado $|\psi\rangle$, esto lo podemos hacer dado que estamos asumiendo que el estado es una propiedad que se le asocia a sistemas individuales. Luego, después de una medición de la propiedad Ξ donde obtuvimos un resultado ξ, asociamos el estado $|\xi\rangle$ al sistema.

Es decir, de antes a después de la medición el estado del sistema cambió de la siguiente manera:

$$|\psi\rangle \rightarrow |\xi\rangle \tag{8.2}$$

a este proceso es a lo que normalmente nos referimos como el "colapso".

Esto es algo que seguro le parece a usted familiar. Permítame citar a Dirac en su explicación de esta situación:

> We now make some assumptions for the physical interpretation of the theory. If the dynamical system is in an eigenstate of a real dynamical variable ξ, belonging to the eigenvalue ξ', then a measurement of ξ will certainly give as result the number ξ'. Conversely, if the system is in a state such that a measurement of a real dynamical variable ξ is certain to give one particular result (instead of giving one or other of several possible results according to a probability law, as is in general the case), then the state is an eigenstate of ξ and the result of the measurement is the eigenvalue of ξ to which this eigenstate belongs.

Esta es la primera parte del argumento, si el sistema está en el estado $|\xi\rangle$, al medir obtendremos el valor ξ. Si medimos el valor ξ, sabemos que el sistema está en el valor ξ.

Vamos a insistir en que este texto está escrito de manera que el estado se aplique al sistema individual y sea objetivo. Esto es de relevancia para la primera parte. Se dice que el sistema dará el número ξ cuando está en el estado $|\xi\rangle$. Del fraseo se puede inferir que se refieren a solo un sistema al que se le medirá la propiedad y dará un solo resultado la medición como es de esperarse.

Si quisiéramos extender esta idea al caso donde el estado se aplica solo a grupos de sistemas se necesita modificar ligeramente. Por ejemplo, que si para un conjunto de sistemas, al medir la propiedad ξ todos dan el valor ξ' podemos decir que estado del conjunto es $|\xi\rangle$. De la misma manera, si un conjunto está en el estado $|\xi\rangle$, al medir la propiedad ξ a cualquiera de los elementos obtendremos el valor ξ.

Parece un poco trivial gastar tiempo en estas definiciones que parecen naturales hasta cierto punto. Pero pongamos ahora atención a la parte donde Dirac menciona específicamente el salto que pasa en la evolución:

> When we measure a real dynamical variable ξ, the disturbance involved in the act of measurement causes a jump in the state of the dynamical system. From physical continuity, if we make a second measurement of the same dynamical variable ξ immediately after the first, the result of the second measurement must be the same as that of the first. Thus after the first measurement has been made, there is no indeterminacy in the result of the second. Hence, after the first measurement has been made, the system is in an eigenstate of the dynamical variable ξ, the eigenvalue it belongs to begin equal to the result of the first measurement. This conclusion must still hold if the second measurement is not actually made. In this way we see that a measurement always causes the system to jump into an eigenstate of the dynamical variable that is being measured, the eigenvalue this eigenstate belongs to being equal to the result of the measurement.

Nótese primero que se acepta que el sistema tenga un salto en el estado debido a una medición. Luego supone que al medir inmediatamente se debe repetir ese valor. Luego, según lo citado antes, podemos decir cuál es el estado del sistema. [1]

Ahora pensemos en este argumento en el caso que se quiera tener un estado cuántico que se aplique a grupos. Mediremos una propiedad al grupo y para cada medición tendremos un resultado, en general diferente. Sin embargo, no podremos asociar un estado cuántico al conjunto en total pues nuestra regla exige que todas las mediciones resulten en el mismo valor. Luego nos quedamos sin herramientas para asociarle un estado a este grupo después de la medición.

Sin embargo, es cierto que si del grupo original separamos todos los sistemas que tienen el mismo valor, tendremos un nuevo grupo que sí cumple las suposiciones de nuestra regla. Así que sí podremos asociarle el eigenestado correspondiente al valor.

Lo importante de lo anterior es que no hubo un cambio en el estado cuántico, pues se asocia a grupos. Lo que cambió es el grupo. Primero un grupo tenía un estado cuántico, luego hicimos otro grupo que tiene asociado otro estado cuántico. Es decir no hubo un cambio del estado sino un cambio en el grupo al que le estamos poniendo atención.

[1] Nótese cómo se evoca a la continuidad para justificar el paso crucial. Esto de hecho no estaba motivado por la visión de la mecánica cuántica del momento. Había cierto consenso en lo que se refiere a que la medición afectaba al sistema. Sin embargo, dicha afectación aplica para cualquier medición. Véase Bohr y Heisenberg

También hay otra consecuencia importante que viene de exigir que el estado se asocie a grupos. Tenemos una regla para saber cómo asociar un estado cuando todos los elementos del grupo dan el mismo valor en una medición. Pero no tenemos ninguna garantía de que si esto no pasa vamos a tener algún estado cuántico que pueda describir a ese grupo. Bien podría ser el caso de que nos encontremos con un grupo para el que no podamos encontrar un estado cuántico que describa el resultado de sus mediciones.

En contraposición, en la visión del estado cuántico que se puede aplicar a sistemas individuales, siempre podemos asociar un estado cuántico a un sistema que acabamos de medir. Ese es todo el punto de la segunda cita de Dirac.

Para concluir me gustaría de nuevo insistir en que el proceso del colapso, como normalmente lo entendemos, necesita forzosamente de que el estado cuántico se aplique a sistemas individuales. Así, además, el estado cuántico se puede aplicar a cualquier sistema. Tenemos garantizado que podremos asociar un estado cuántico a cualquier sistema.

En cambio, una visión del estado cuántico que solo aplique a grupos, no podremos describir un proceso similar. Sin embargo, tampoco tendremos la garantía de que se puede asociar un estado cuántico a cualquier grupo de sistemas.

No quisiera terminar esta sección sin decir algunas palabras la imagen pública de este problema. La discusión sobre si el estado cuántico es una cantidad que se asocia a un sistema individual o a un grupo de sistemas es poco discutida pues me parece se confunde o se mezcla con la famosa discusión sobre la completitud de la mecánica cuántica.

Me parece necesario decir que el término completitud es un abuso horrible de lenguaje. Si una teoría es completa nunca ha sido un criterio para una buena teoría física. Es más por qué nos interesaría que una teoría sea completa, ¿no podemos aspirar a tener teorías que solo describan lo que pueden en su rango de aplicación y no mucho más? Además de esto, la batalla pública entre lo que ahora mal llamamos la escuela de Copenhagen y los que mantenían una visión más clásica, dejó divisiones marcadas. Sumado a la decadencia de la profesión científica que vino con la segunda guerra mundial y la superstición científica que viene en la segunda mitad del siglo XX; esta cuestión se ha vuelto una tierra seca e infértil.

Sobre ella solo me gustaría mencionar que incluso los personajes centrales de la cuestión tenían ideas que no corresponden a la imagen pública que hoy tenemos. La principal es probablemente la visión de Einstein sobre el indeterminismo. El mejor lugar para verificar la visión de Einstein es, contrario a todo sentido común, en los escritos de Einstein. Particularmente en las cartas Einstein-Born. Ya al final de la vida de ambos, fuera de los debates centrales de la ciencia del momento, encontramos algunos pasajes como este:

> The other remark concerns your interpretation of the ψ-function; it seems to me that it completely agrees with what I have been thinking all along, and what most reasonable physicists are thinking today. To say that ψ describes the 'state' of one single system is just a figure of speech, just as one might say in everyday life: 'My life expectation (at 67) is 4.3 years'. This, too, is a statement about one single system, but does not make sense empirically. For what is really meant is, of course, that you take all individuals of 67 and count the percentage of those who live for a certain length of time. This has always been my own concept of how to interpret $|\psi|^2$. Instead you propose a system of a large number of identical individuals - a statistical total. It seems to me that the difference is not essential, but merely a matter of language.

Esta cita me parece importante pues revela la visión de Born sobre probabilidad. Una visión cercana al frecuentismo sin más, esto de esperarse pues los círculos de Born son también los círculos de von Mises[33].

Einstein, en su respuesta, responde coincidiendo con Born en lo que se refiere a ensambles [2] y acepta que, para él, la característica que hace a la teoría, o descripción, ser incompleta es justo que no es una descripción del sistema individual[24, pág. 186]:

> I see from the last paragraph of your letter that you, too, take the quantum theoretical description as incomplete (referring to an ensemble). But you are after all convinced that no (complete) laws exist for a complete description, according to the positivistic maxim *esse est percipi*. Well, this is a programmatic attitude, not knowledge. This is where our attitudes really differ.

Esto nos lleva a concluir que al menos Einstein y Born solo aceptaban que el estado cuántico se podía asociar a grupos. Diferirán, claro de Bohr y otros de la escuela de Copenhagen[56], en este sentido. Esto conecta con

[2]Nótese también que al referirse a ensambles en este contexto es completamente diferente al de la mecánica estadística. Ahí el ensamble es una técnica de cálculo que no tiene correspondiente en el mundo físico. Aquí ensamble se refiere a un grupo grande de sistemas o eventos.

nuestra discusión anterior pues Einstein entendía claramente que asociar el estado cuántico a una descripción individual tenía problemas como el colapso que discutimos anteriormente y una interpretación donde el estado se asocia a grupos no pasa por este problema. Para más sobre la visión de Einstein véase [8].

8.2. De lo que llaman colapso

Vamos a discutir la famosa cuestión del colapso usando algunos de los elementos que hablamos en la primera parte.

Supongamos que en la regla de Born queremos usar una probabilidad epistémica, ya sea clásica, logicista o subjetiva; pero queremos que el estado cuántico sea una propiedad objetiva. Este caso es de mucha importancia para nuestra discusión.

Imaginemos que en cierta situación la ecuación de Born se cumple. La probabilidad, al ser epistémica depende del conocimiento o pensamiento de la persona que la asigna. Llamemos a este conocimiento E, así que la probabilidad le pondremos un subíndice para denotar esta dependencia $P_E(X = x)$.

Así la ecuación 8.1 toma la forma:

$$|\langle x \,|\, \psi \rangle|^2 = P_E(X = x). \tag{8.3}$$

Esto ya nos presenta un problema, si dos personas tienen diferentes conocimientos, digamos E y E', en general asignarán distintas probabilidades,

$$P_E(X = x) \neq P_{E'}(X = x). \tag{8.4}$$

la ecuación anterior simplemente no se puede sostener. Si el estado cuántico es objetivo, su valor tiene que ser el mismo para ambas personas. Luego las siguientes ecuaciones no se pueden sostener al mismo tiempo y ser consistentes con 8.4:

$$|\langle x \,|\, \psi \rangle|^2 = P_E(X = x). \tag{8.5}$$

$$|\langle x \,|\, \psi \rangle|^2 = P_{E'}(X = x). \tag{8.6}$$

Esto es de hecho algo que parece trivial, pero es fácil pasarlo por alto. Utilizar de un lado de la ecuación una propiedad objetiva y del otro una epistémica es insostenible.

Esta situación puede parecer familiar cuando analizamos la interpretación a veces llamada como de von Neumann-Wigner. Esta interpretación dice que el estado cuántico colapsa cuando la persona que realiza la medición se entera de su resultado. A veces se dice que es la conciencia de quien realiza la medición la que actúa como aparato de medición y termina por colapsar el estado cuántico.

No es difícil ver que estas aseveraciones se pueden entender dentro del marco que acabamos de discutir. Cuando se intenta mantener un estado cuántico objetivo, pero una probabilidad clásica, la regla de Born tiene problemas. El lado de la probabilidad cambiará cada vez que haya alguna nueva evidencia y la única forma de mantener la regla será que el lado izquierdo tenga cambios los que no tendrán una explicación objetiva. Seguro a usted le es familiar esto.

Esto no es un problema especial de la mecánica cuántica. La física ha tenido este problema en particular desde la introducción de la probabilidad, digamos desde los primeros trabajos más bien conjeturales de Bernoulli que reinterpretaban variables como la presión en términos de variables mecánicas microscópicas como el momento.

Para muestra basta un botón. Digamos la distribución de velocidades en un gas. Estamos acostumbrados a poner que la velocidad es:

$$V = \sum_i v_i p_i \tag{8.7}$$

de nuevo tenemos de lado izquierdo una propiedad que se asume como objetiva, pero de lado derecho tenemos la probabilidad que es epistémica y permite que el lado derecho sea diferente si lo aplican diferentes personas.

Cuántas ecuaciones tenemos en la física que justo tienen este problema. De un lado hay una propiedad objetiva, pero del otro lado tenemos un valor promedio, desviación estándar u otra variable que depende de probabilidades. Todas esas ecuaciones tienen este mismo problema, no podemos sostener que un lado de la ecuación sea objetivo y otro sea epistémico.

Es cierto que este hecho es muy simple como para dedicarle tanto tiempo, pero realmente es crucial en el entender nuestra situación.

No es difícil pensar que en un intento terco por sostener la ecuación 8.1, un estado objetivo y una probabilidad epistémica, pueda sostener un discurso donde al cambiar la probabilidad, digamos porque su conocimiento cambio gracias a que vio el resultado de una medición, se afirme que el estado cambió. No esperamos que esto pase pues se afirma que el estado es objetivo.

Naturalmente, hablo de una de las formas en las que el término “colapso” es usado hoy día y que ha sido criticada constantemente, pero sigue en buen uso por una parte de la comunidad.

Le invito a revisar todas las posibilidades y pensar si ha leído una argumentación que use más de una. Es más, le invito a revisar si en sus propias argumentaciones alguna vez ha usado más de una. La respuesta no es de sorprender. Es a esta falta de coherencia con la probabilidad a la que nos hemos referido antes. No es una cuestión sólo de la mecánica cuántica, toda la ciencia sin distinción abusa del término probabilidad. Utilizamos una visión inconsistente de la probabilidad que parece que nos permite abordar cualquier problema que el mundo nos ponga enfrente. Saltamos entre una interpretación y otra de la probabilidad con la facilidad que una ardilla salta entre las ramas de varios arboles y lo hacemos con la misma naturalidad que estas criaturas. Tal es la práctica que tenemos haciendo esto que realmente es difícil ver cuando lo hacemos o dejar de hacerlo.

Capítulo 9

Las muchas desigualdades de Bell

Ya hemos criticado la aplicación de los axiomas de Kolmogorov sin una escuela de pensamiento que respalde el uso de los axiomas. Aplicar los axiomas de Kolmogorov sin un pensamiento sobre la correcta aplicación de la teoría permite demasiadas posibilidades que son inconsistentes entre sí. Para demostrar hasta donde es posible llevar esto analizaremos ahora las desigualdades de Bell. Me gustaría subrayar que el siguiente análisis se hace estrictamente dentro de los axiomas de Kolmogorov; no agregaremos ninguna de las escuelas de pensamiento discutidas antes.

Al estar estrictamente dentro del terreno de los axiomas de Kolmogorov, vamos a seguir estrictamente sus reglas. No será esto una clásica descripción de libro de texto sobre las desigualdades, donde saltaremos a valores esperados inmediatamente. Con "estrictamente" nos referimos a que tendremos que discutir las precondiciones de los axiomas, es decir el espacio muestral y el espacio de eventos(el axioma I y II originales perdidos en la decadencia de la probabilidad) y después podremos pasar a escribir las desigualdades. Esto nos llevará a diferentes resultados.

Normalmente nos referimos a las desigualdades de Bell en plural. Hay de hecho muchas formas de derivar estas desigualdades. Sin embargo, la mayoría pensamos que la desigualdad a la que llega cada demostración es la misma. Es decir, pensamos que las desigualdades a las que se llega son todas equivalentes matemáticamente. Se puede pasar de una a otra con transformaciones matemáticas que, se asume, no cambian el significado del resultado.

Es una sorpresa para muchos de nosotros que lo anterior no es correcto.

Hay de hecho varios objetos que se nombran como desigualdades de Bell y no son equivalentes entre sí. No son equivalentes en el sentido matemático, pues no son condiciones necesarias y suficientes entre sí, y no son equivalentes en significado. Aquí describiremos al menos dos[72].

9.1. El esquema de experimentos EPRB tiene una descripción en al menos dos espacios de probabilidad distintos

En los experimentos del tipo EPRB se asume que un sistema físico es divido en dos partes. Cada parte se manda a un arreglo experimental distinto. En cada aparato hay un ángulo que puede tomar valores a disposición del operador. Este arreglo experimental dará una cantidad que solo puede tomar uno de dos valores, normalmente con valores $+1$ y -1.

En el primer caso tendremos un par de variables aleatorias para cada lado del experimento, digamos (A, α) y (B, β). Agregamos además una posible variable Λ que es común para ambos lados del experimento. De acuerdo a nuestra exposición anterior, un uso de la probabilidad en los estándares actuales requiere un espacio muestral, uno de eventos y una medida de probabilidad.

En este caso el espacio muestral estará compuesto por quintetas de la forma $(A = a, B = b, \alpha = \theta, \beta = \theta', \Lambda = \lambda)$, recordando que a, b son binarias 0 o 1, θ, θ' son ángulos, que por simplicidad sólo les permitiremos tener dos valores 0 y 45; y λ pertenece a un conjunto no especificado.

Pongamos por ejemplo que en el experimento ponemos $\alpha = 0$, $\beta = 45$ y que obtenemos $A = 1$ y $B = 0$. No sabemos qué valor tuvo Λ. Dentro del espacio de eventos del espacio muestral anterior, el evento que se le asocia a esta situación es:

$$\begin{aligned} \Big\{ &(A = 1, \alpha = 0, B = 0, \beta = 45, \Lambda = \lambda'), \\ &(A = 1, \alpha = 0, B = 0, \beta = 45, \Lambda = \lambda''), \\ &\qquad\vdots \\ &(A = 1, \alpha = 0, B = 0, \beta = 45, \Lambda = \lambda^{(n)}) \Big\} \end{aligned} \tag{9.1}$$

donde tenemos una cuarteta por cada posible valor para Λ. Naturalmente si el valor de lambda fuera conocido tendríamos solo una quinteta en este evento. Esta descripción es correcta bajo todos los criterios de la probabilidad de Kolmogorov que describimos anteriormente.

Ahora pongamos otra descripción de la situación. De nuevo asumimos, por simplicidad, que los ángulos solo pueden tener dos valores. Para cada lado del experimento elegimos dos variables aleatorias A_0 y A_{45}, y las análogas para el lado B. La variable A_0 es el resultado de la medición si el ángulo α tuviera el valor 0; la variable A_{45} será el resultado de la medición si el ángulo α tuviera el valor 45. Al igual que el caso anterior agregaremos una variable Λ para representar cualquier otro parámetro.

El espacio muestral en este caso estará hecho de quintetas de del siguiente tipo:

$$(A_0 = a, A_{45} = a', B_0 = b, B_{45} = b', \Lambda = \lambda) \tag{9.2}$$

Si volvemos a la situación descrita para el espacio muestral 1, tenemos que el evento que le corresponde en este espacio muestral es diferente:

$$(A_0 = a, A_{45} = a', B_0 = b, B_{45} = b', \Lambda = \lambda) \tag{9.3}$$

$$\begin{Bmatrix} (A_0 = 1, A_{45} = 0, B_0 = 0, B_{45} = 0, \Lambda = \lambda), \\ (A_0 = 1, A_{45} = 1, B_0 = 0, B_{45} = 0, \Lambda = \lambda), \\ (A_0 = 1, A_{45} = 0, B_0 = 1, B_{45} = 0, \Lambda = \lambda), \\ (A_0 = 1, A_{45} = 1, B_0 = 1, B_{45} = 0, \Lambda = \lambda), \\ \vdots \end{Bmatrix} \tag{9.4}$$

donde tenemos estas cuatro quintetas para cada valor posible de Λ.

De nuevo queremos resaltar que esta descripción es completamente correcta desde el punto de vista de la probabilidad de Kolmogorov como fue expuesta anteriormente.

9.2. Una desigualdad formalmente igual a la de Bell se puede construir en cada espacio de probabilidad expuesto

Hoy estamos acostumbrados a ver la desigualdad de Bell en su formulación CHSH que se expresa en términos de valores esperados de la siguiente manera:

$$-2 \leq \langle XY \rangle + \langle X'Y \rangle + \langle XY' \rangle - \langle X'Y' \rangle \leq 2 \tag{9.5}$$

donde cada variable solo puede tomar valores $+1$ y -1. Aquí no aparece la probabilidad directamente. De hecho podría argumentarse que la probabilidad no juega ningún papel pues los valores esperados se pueden comparar directamente con promedios en mediciones del experimento. Más adelante demostraremos que en realidad los supuestos para mantener esta desigualdad serán equivalentes a la probabilidad.

En las deducciones de libro de texto de la desigualdad de Bell se pueden encontrar a dichos promedios expresados de la siguiente manera:

$$\langle XY \rangle = \int X(\alpha, \lambda) Y(\beta, \lambda) \rho(\lambda) d\lambda \tag{9.6}$$

Esta expresión no determina el espacio de probabilidad a donde la distribución ρ pertenece. Pareciera que α, β no actúan como variables aleatorias sino como constantes pero no es definitivo. Con el fin de revelar mejor las relaciones entre variables y establecer claramente cuáles son aleatorias y cuáles parámetros o constantes, pongamos la misma expresión utilizando solo probabilidades. Para esto usaremos unas nuevas variables A, B donde $A = 1$ si $X = +1$ y $A = 0$ si $X = -1$. Esto convierte los promedios en:

$$\begin{aligned} \langle AB \rangle =& 4P(A = 1, B = 1 | \alpha = \theta) \\ & - 2P(A = 1 | \alpha = \theta) - 2P(B = 1 | \beta = \theta) + 1 \end{aligned} \tag{9.7}$$

De acuerdo con lo anterior, podemos ver que esta expresión se encuentra representada en nuestra segunda opción para el espacio muestral. Es decir, ese con solo cinco variables aleatorias: A, B, α, β y Λ. Esto implica que el *espacio muestral* que se utiliza es de la forma $\{(A = a, B = b, \theta = a, \theta' = b, \Lambda = \lambda)\}$.

Con esto se puede cambiar la forma de las desigualdades de Bell; es decir, la ecuación 9.5 se transforma en:

$$\begin{aligned}-1 \leq P(A=1,B=1|\alpha=\theta,\beta=\omega)+P(A=1,B=1|\alpha=\theta',\beta=\omega)\\ +P(A=1,B=1|\alpha=\theta,\beta=\omega')-P(A=1,B=1|\alpha=\theta',\beta=\omega')\\ -P(A=1|\alpha=\theta)-P(B=1|\beta=\omega)\leq 0. \end{aligned}\tag{9.8}$$

La equivalencia matemática se puede consultar en el apéndice A. Donde podemos apreciar claramente cuáles variables están en el condicional de cada término. Además podemos asegurar que el espacio muestral que se usa para estas probabilidades es el de las quintetas que mencionamos antes.

Ahora pasemos a describir la misma desigualdad en el segundo espacio muestral que presentamos. Tendremos entonces las siguientes variables aleatorias: A_0, A_{45}, B_0, B_{45} y las posibles variables ocultas que de nuevo se representan por λ. En este espacio la desigualdad toma una diferente:

$$\begin{aligned}-1 \leq P(A_0=1,B_0=1)+P(A_0=1,B_{45}=1)\\ +P(A_{45}=1,B_0=1)-P(A_{45}=1,B_{45}=1)\\ -P(A_0=1)-P(B_0=1)\leq 0. \end{aligned}\tag{9.9}$$

Notemos la analogía entre las dos desigualdades. Parece que solo hemos cambiado las condicionales por subíndices. Sin embargo ya hemos explicado los distintos espacios muestrales posibles.

9.3. La demostración de las desigualdades utiliza hipótesis distintas en cada espacio.

El objetivo de esta sección es demostrar ambas desigualdades. Como se verá dichas demostraciones requieren hipótesis diferentes. Hay demostraciones mucho más ilustrativas, pero estamos interesados en resaltar las hipótesis involucradas.

9.3.1. Primera desigualdad

Para demostrar la primera desigualdad, es decir, la ecuación 9.8 empecemos por la desigualdad matemática

$$-1 \leq ab+ab'+a'b-a'b'-a-b \leq 0$$

que es válida para cualesquiera números $0 \leq a, a', b, b' \leq 1$. Vamos a realizar la sustitución $a = P(A = 1|\alpha = \theta, \Lambda = \lambda)$ y las expresiones análogas para obtener:

$$\begin{aligned} -1 &\leq \\ -P(A = 1|\alpha = \theta, \Lambda = \lambda)P(B = 1|\beta = \omega, \Lambda = \lambda) \\ -P(A = 1|\alpha = \theta', \Lambda = \lambda)P(B = 1|\beta = \omega, \Lambda = \lambda) \\ -P(A = 1|\alpha = \theta, \Lambda = \lambda)P(B = 1|\beta = \omega', \Lambda = \lambda) \\ +P(A = 1|\alpha = \theta', \Lambda = \lambda)P(B = 1|\beta = \omega', \Lambda = \lambda) \\ +P(A = 1|\alpha = \theta, \Lambda = \lambda) \\ +P(B = 1|\beta = \omega, \Lambda = \lambda) \\ \leq 0. \end{aligned} \tag{9.10}$$

Para continuar pedimos la siguiente hipótesis:

$$P(A = 1|\alpha = \theta, \Lambda = \lambda)P(B = 1|\beta = \omega, \Lambda = \lambda) = P(A = 1, B = 1|\alpha = \theta, \beta = \omega, \Lambda = \lambda)$$

que aquí le llamaremos *hipótesis de separabilidad*, pero se puede encontrar en la literatura con otros nombres como *hipótesis de localidad* o de *factorizabilidad*. [10]

Esto nos permite llegar a la expresión:

$$\begin{aligned} -1 &\leq \\ P(A = 1, B = 1|\alpha = \theta, \beta = \omega, \Lambda = \lambda) \\ +P(A = 1, B = 1|\alpha = \theta', \beta = \omega, \Lambda = \lambda) \\ +P(A = 1, B = 1|\alpha = \theta, \beta = \omega', \Lambda = \lambda) \\ -P(A = 1, B = 1|\alpha = \theta', \beta = \omega', \Lambda = \lambda) \\ -P(A = 1|\alpha = \theta, \Lambda = \lambda) \\ -P(B = 1|\beta = \omega, \Lambda = \lambda) \\ \leq 0. \end{aligned} \tag{9.11}$$

El siguiente paso consiste en hacer un promedio sobre los posibles valores λ para eliminarla de la desigualdad. Por ejemplo, para el último término deseamos hacer uso de:

$$\int P(B = 1|\beta = \omega, \Lambda = \lambda)P(\Lambda = \lambda|\beta = \omega)d\lambda = P(B = 1|\beta = \omega)$$

Sin embargo cada término de la desigualdad tiene α y β con valores distintos. Luego se pide la siguiente hipótesis:

$$P(\Lambda = \lambda|\alpha, \beta) = P(\Lambda = \lambda)$$

A esta hipótesis se le llamaremos *independencia de* λ y también se puede encontrar con otros nombres como *no-contextualidad*.

La hipótesis anterior nos permite multiplicar toda la desigualdad por $P(\lambda)$ y realizar una suma, integral o el procedimiento para promediar que sea necesario con la seguridad de que la variable Λ será eliminada. Esto nos lleva a la expresión:

$$\begin{aligned} -1 \leq \int \Big(& P(A=1, B=1|\alpha=\theta, \beta=\omega, \Lambda=\lambda) \\ & +P(A=1, B=1|\alpha=\theta', \beta=\omega, \Lambda=\lambda) \\ & +P(A=1, B=1|\alpha=\theta, \beta=\omega', \Lambda=\lambda) \\ & -P(A=1, B=1|\alpha=\theta', \beta=\omega', \Lambda=\lambda) \\ & -P(A=1|\alpha=\theta, \Lambda=\lambda) \\ & -P(B=1|\beta=\omega, \Lambda=\lambda)\Big) P(\Lambda=\lambda) d\lambda \leq 0. \end{aligned} \tag{9.12}$$

Esta expresión es igual a la primera desigualdad de Bell que mencionamos, ec. 9.8.

9.3.2. Segunda desigualdad

Ahora pasemos a la demostración de la desigualdad en el segundo espacio muestral.

Empezamos de nuevo con la desigualdad

$$-1 \leq ab + ab' + a'b - a'b' - a - b \leq 0$$

que nos permite realizar la asignación:

$$a = A_0, a' = A_{45}, b = B_0, b = B_{45}$$

pues cada variable solo toma valores 0 o 1.

Multiplicamos la desigualdad por $P(A_0, A_{45}, B_0, B_{45})$ para obtener:

$$\begin{aligned} -P(A_0, A_{45}, B_0, B_{45}, \lambda) \leq \{ & A_0B_0 + A_{45}B_0 + A_0B_{45} - A_{45}B_{45} \\ & -A_0 - B_0\} P(A_0, A_{45}, B_0, B_{45}, \lambda) \leq 0 \end{aligned} \tag{9.13}$$

Ahora realizamos un promedio sobre todas las variables:

$$\begin{aligned} -1 \leq \int_\Lambda \sum_{A_0, B_0, A_{45}, B_{45} = \{0,1\}} \{ & A_0B_0 + A_{45}B_0 + A_0B_{45} - A_{45}B_{45} \\ & -A_0 - B_0\} P(A_0, A_{45}, B_0, B_{45}, \lambda)\, d\lambda \leq 0 \end{aligned} \tag{9.14}$$

Cuadro 9.1: Hipótesis necesarias para la demostración de las desigualdades

	Desigualdad 1	Desigualdad 2
Hipótesis Requeridas	Espacio de probabilidad 1 Axiomas de Kolmogorov Localidad λ-independencia	Espacio de probabilidad 2 Axiomas de Kolmogorov

Debido a que nuestras variables solo tienen valores $0, 1$, tenemos que el valor esperado de una variable es igual a la probabilidad de tener el valor 1, es decir, $< A >= P(A = 1)$. Esto nos lleva a que la mayoría de los términos de la desigualdad se pueden escribir de la siguiente forma:

$$\int_\Lambda \sum_{A_0,B_0,A_{45},B_{45}=\{0,1\}} A_0B_0 \quad P(A_0, A_{45}, B_0, B_{45}, \lambda)\, d\lambda = P(A_0 = 1, B_0 = 1). \tag{9.15}$$

Usando esto en la desigualdad llegamos a la forma deseada, ec. 9.9:

$$\begin{aligned} -1 \leq P(A_0 = 1, B_0 = 1) + P(A_0 = 1, B_{45} = 1) \\ +P(A_{45} = 1, B_0 = 1) - P(A_{45} = 1, B_{45} = 1) \\ -P(A_0 = 1) - P(B_0 = 1) \leq 0. \end{aligned} \tag{9.16}$$

Es de crucial importancia para nuestra discusión que este cálculo no usó hipótesis alguna, solo se usaron propiedades de la probabilidad, léase los axiomas de Kolmogorov, para las cuatro variables aleatorias A_0, B_0, A_{45}, B_{45} y el hecho de que éstas tienen valores 0 o 1. Hemos llegado así a una desigualdad de la misma forma que la de Bell.

9.4. Hay, al menos, dos desigualdades de Bell

Hasta ahora planteamos dos espacios de probabilidad para las mediciones de espín. Desarrollamos dos desigualdades formalmente iguales, una en cada espacio de probabilidad. Finalmente demostramos dicha desigualdad utilizando diferentes juegos de hipótesis.

La desigualdad del espacio uno requiere de la validez de los axiomas de Kolmogorov, la hipótesis de localidad y la hipótesis de λ-independencia. La desigualdad del espacio dos solo requiere los axiomas de Kolmogorov. Continuemos sobre esta cuestión.

9.5. Hay distintas posibilidades para interpretar la violación de las desigualdades de Bell en experimentos cuánticos

Desde hace varios años hay realizaciones de experimentos del tipo EPRB[40]. En todos ellos se ha violado dicha desigualdad. Se les ha celebrado, confirmado, analizado, criticado y hasta negado. Se han propuesto muchas formas de analizarlos y cosas que pudieron salir mal. Sin abordar la validez de dichos experimentos, exploremos desde nuestra mirada qué nos dice una violación de las desigualdades.

El primer espacio donde se formula la desigualdad descansa en tres hipótesis como ya se dijo. Si aceptamos que la desigualdad se viola debemos concluir que al menos una de las tres hipótesis no es válida para el experimento. De nuevo, hay una extensa literatura sobre la hipótesis de separabilidad y la de contextualidad. Se habla de efectos no locales, contextuales, realismo, etc. Este no es nuestro objetivo.

Desde el segundo espacio donde obtuvimos la desigualdad solo se tienen los axiomas de Kolmogorov. Si la desigualdad no se verifica, necesariamente los axiomas no son válidos para nuestra situación.

En el trabajo original de Fine[29] donde se obtuvo esta versión de la desigualdad, se propone que la hipótesis directamente responsable es la existencia de la probabilidad conjunta $P(A_0, B_0, A_{45}, B_{45})$. Probablemente pasó desapercibido, pero es necesario asumir que existe para obtener la desigualdad en el segundo espacio[1].

No olvidemos que negar la existencia de la probabilidad conjunta va directamente en contra de Kolmogorov. Ya hablamos en la primera parte que la probabilidad debe estar definida en cada elemento del espacio de eventos. Luego, negar que la probabilidad conjunta existe, es negar los requerimientos para establecer los axiomas de Kolmogorov.

Hay otras posibilidades, por ejemplo, en vez de negar los requerimientos de los axiomas de Kolmogorov, podemos negar la validez de los axiomas en sí. También podemos encontrar mucho trabajo al respecto [61]. No es difícil encontrar personas que aseguran que la probabilidad que describe a los fenómenos cuánticos es de distinta naturaleza que la probabilidad cotidiana. Incluso se puede encontrar términos como la *probabilidad cuántica.* Esta posibilidad será estudiada parcialmente, sin embargo, no es nuestra intención cambiar los axiomas de Kolmogorov para describir los fenómenos cuánticos.

[1]También en el primero hay que suponer que ciertas distribuciones existen, pero no es de interés en esa comunidad.

Como ya se dijo en la primera parte, no nos parece adecuado que se use la axiomática de Kolmogorov para describir *la probabilidad.* Por un lado, una mínima visión de esta disciplina revela que no existe **un** solo concepto de probabilidad. Por otro, se puede argumentar que los axiomas de Kolmogorov no describen correctamente ninguna de las escuelas de pensamiento sobre la probabilidad.

Además de todas las posibilidades argumentales, la cuestión importante parece ser cuál de las dos aproximaciones a los experimentos EPRB es la *correcta.* Cuál de los espacios de probabilidad que expusimos es el que *se debe* usar para estos experimentos. Después de todo parece que de esto dependen elementos importantes para la física actual, como la no localidad.

Nos parece que esta cuestión cambiará según la teoría de la probabilidad que se use. No se puede pensar que la misma argumentación servirá para todas. Además el discurso que acompaña a las distintas teorías será determinante para argumentar en favor o en contra de cada opción.

Capítulo 10

Aún más desigualdades de Bell

En el capítulo pasado intentamos dar cuenta de las posibilidades de expresar un experimento tipo Bell usando los axiomas de Kolmogorov. Escribimos un par de alternativas de las muchas que hay. Espero que esto sea suficiente para convencerle de que los axiomas de Kolmogorov son solo una teoría matemática, un conjunto de axiomas. Por sí mismos no son capaces de determinar la aplicación de la probabilidad a ejemplos concretos. Luego no debe sorprendernos que nos encontremos con diferentes formas de aplicarlos que matemáticamente se vean iguales, pero lleven a conclusiones diferentes.

¿Qué pasa si ahora intentamos seguir el argumento de las desigualdades en un contexto de probabilidad más estricto? ¿Podremos seguir formulando estas dos versiones de la desigualdad y llegaremos a las mismas conclusiones?

Tres puntos fundamentales para el argumento de las desigualdades son:

- Las hipótesis de separabilidad y no-contextualidad
- El espacio de probabilidad y desigualdad que usan
- La verificación experimental de la violación

Son las primeras dos partes las que interesan para este trabajo. La tercera nos parece que nunca ha tenido verdaderos retos y los “loopholes” que han sido tan mencionados en la literatura solo han sido una forma de simular una carrera por cerrarlos[1] cuando ya nadie esperaba que hubiera ningún resultado diferente.

[1] Excelente forma de generar atención pública ante una cuestión científica que está muerta.

10.1. La hipótesis de separabilidad

Si vemos las desigualdades desde un punto de vista frecuentista, hay un par de consideraciones que debemos tener para poder seguir el argumento.

Primero, es que la hipótesis de separabilidad no tiene una relación con la localidad o la causalidad tan fuerte como normalmente se sugiere en los textos. Por ejemplo, en *The theory of local beables*[10] podemos encontrar una motivación de la hipótesis de separabilidad

Es justo la conexión entre la causalidad y las probabilidades condicionales la que justifica que esta hipótesis se le pueda dar alguna relación con la localidad. Desafortunadamente, dicha relación intuitiva está más en el terreno de la probabilidad epistémica y viene de los tiempos de la probabilidad clásica.

En el caso del frecuentismo, hay un problema importante pues el frecuentismo no tiene tan buena relación con la causalidad. De hecho, von Mises se muestra sumamente crítico con el concepto de causalidad[55, pág. 210]:

> Using the tools of probability and statistics, these new concepts have led us from the kinetic theory of gases to quantum theory and wave mechanics. It now appears inevitable that we must abandon another cherished notion that has its origin in everyday life and pre-scientific thought and has been elevated to the rank of an eternal category of thought by overly zealous philosophers : the naive concept of causality.

Von Mises continua esta sección con un par de ejemplos. Pensemos que tenemos un dado que lanzamos muchas veces para tener un colectivo y sus resultados favorecen al número tres, es decir, la probabilidad en ese colectivo es mayor para tres. Hasta ahí todo normal. Luego si queremos explicar dicho sesgo será natural que digamos, por ejemplo, que el centro de masa del dado no está en el centro. Dicha explicación, dice von Mises, depende de la física del momento. Cita él el ejemplo de como toda la noción de causalidad se vió modificada para dar cuenta de la ley de inercia galilieana. Así, la crítica de von Mises a la causalidad es que ésta es un concepto secundario que cambia según el momento.

Este ejemplo nos permite introducir las propensiones en esta discusión. En el ejemplo del dado, decimos que la causa del sesgo es el centro de masa. Lo interesante es que nombramos a las condiciones del lanzamiento como las causas del sesgo. Esta es una de las razones que dieron fuerza a la teoría de las propensiones.

En general, las causas que se pueden dar para un sesgo o cualquier otra propiedad de un colectivo serán condiciones que comparten los elementos del colectivo. Esto es normal, si ningún elemento del colectivo compartiera alguna característica sería increíblemente difícil dar una causa, pues no habría causas compartidas entre sus elementos. Luego la relación a la que estamos acostumbrados entre la causalidad y la probabilidad encuentra una mejor fundamentación en la teoría de las propensiones.

Ya sea propensiones de eventos individuales o propensiones de grupo, en ambos casos son las condiciones que definen la propensión las que se pueden considerar como las causas de los sesgos, promedios, desviaciones o cualquier comportamiento que se observe.

De hecho tanto Popper como Fetzer sugirieron en su momento que las propensiones se podían considerar una forma más relajada de causalidad[36, pág. 129], aunque con algunos problemas para lograr cerrar dicha interpretación, e.g. [37].

Teniendo en cuenta lo anterior y sumado al hecho de que las propensiones siempre se consideran probabilidades condicionadas a, valga la redundancia, las condiciones en las que los eventos se llevan a cabo, tenemos una excelente fundamentación para que la hipótesis de separabilidad se relacione sin mayor problema con la causalidad, tal y como *The theory of local beables* intenta hacer[76]

Ahora también podemos agregar si se está hablando de una propensión de evento individual o de largo plazo. En una buena parte de los escritos de Bell podemos ver que se refiere a probabilidades de cada evento, lo que sugiere que para acercarnos a su intención original debemos pensar en una propensión de evento individual.

10.2. Los espacios de probabilidad

Los dos espacios de probabilidad que se expusieron antes se pueden plantear en la teoría frecuentista y en las propensiones. Sin embargo, hay un poco que decir sobre el caso de las propensiones de evento individual.

Dadas unas condiciones S en las que se lleva a cabo un evento E, en algunos libros se representa $P(E|S)$ como la propensión de que el evento E pase dadas las condiciones S. La naturaleza de las propensiones las lleva a ser siempre probabilidades condicionales bajo algún juego de condiciones[36].

Si recordamos, hemos platicado dos desigualdades de Bell. Una tiene términos del tipo

$$P(A=1,B=1|\alpha=\theta,\beta=\omega,\Lambda=\lambda)$$

mientras que la otro tiene términos del tipo:

$$P(A_0=1,B_0=1)$$

El problema con el segundo caso es que no está condicionada a nada. Eso se puede solucionar. Se pueden agregar condicionales a cada término que describan las situaciones en donde el evento se lleva a cabo:

$$P(A_0=1,B_0=1|S_i)$$

.

Sin embargo, cada uno de estos eventos se lleva en condiciones diferentes, en particular, se llevan a cabo para diferentes valores de α y β normalmente interpretados como los ángulos de los detectores.

La parte interesante viene aquí, si estamos en una propensión de caso singular las condiciones S_i tienen que describir completamente la situación. Tienen que incluir los valores α y β y serán diferentes para cada término. Digamos que separamos las condiciones S_i en α y β y el resto de las condiciones las llamamos s_i. Luego los términos de dicha ecuación cambiarían a:

$$P(A_0=1,B_0=1|\alpha,\beta,s_i)$$

, donde α y β asumirían valores distintos dependiendo si nos referimos a A o a A' y lo mismo para B. Es decir, forzosamente tendríamos que trabajar con la otra desigualdad de Bell que ya vimos no es matemáticamente equivalente.

Luego si estamos trabajando con propensiones de caso individual, la desigualdad de Bell de la ecuación 9.9 no se puede formular con los requerimientos que le exige este marco de probabilidades. Solo se puede formular la desigualdad 9.8.

Con esto podemos ver que si tomamos las propensiones de caso individual, además de que la hipótesis de separabilidad tiene la relación con la localidad que esperaríamos, las posibilidades para describir un experimento tipo Bell también se reducen y no hay duda sobre qué espacio de probabilidad debe usarse. Terminamos así con un teorema bien establecido, o tan bien establecido como las propensiones que usemos, donde solo podemos concluir que la hipótesis de separabilidad o la de no contextualidad se violan y ambas se podrán relacionar de alguna manera con la localidad.

Tal vez el único revés de esta descripción sería que las propensiones de caso único tienen una fuerte dependencia de condiciones que bien pueden estar fuera del cono de luz y por tanto ya tienen una componente fuerte de algún tipo de no-localidad, con o sin el teorema de Bell. Debido a que en esta visión la propensión es una probabilidad objetiva, si esta se ve afectada por algo fuera del cono de luz ya podríamos hablar de un efecto no local. Naturalmente hay propensiones que se han limitado para solventar esta situación. Véase [36, 54].

Con respecto a las propensiones de largo plazo y el frecuentismo, ninguno necesita este requerimiento en las condiciones por lo que pueden expresar ambos espacios de probabilidad. Ya no digamos las probabilidades epistémicas que claramente tienen un mayor margen de maniobra al respecto.

Esto nos lleva a una cuestión muy interesante. Digamos que quiero tomar una postura sobre las desigualdades de Bell, digamos además que me gusta el frecuentismo así que tomo esa posición y por otro lado encuentro alguna hipótesis que me permite negar la validez o la pertinencia de la desigualdad de Bell 1, es decir la ecuación 9.8. Así, para mí la discusión sobre la no-localidad no tiene mucho sentido. Para empezar la relación entre la hipótesis de separabilidad y la causalidad no está bien fundada. Luego la desigualdad de Bell 1 tiene argumentos en contra y la forma correcta de describir un experimento tipo Bell es la desigualdad de Bell 2, ecuación 9.9. Dicha desigualdad no necesita la hipótesis de separabilidad. Todo esto me llevaría a la conclusión de que la desigualdad de Bell no tiene nada que ver con la no-localidad. Conclusión: Toda la discusión de no-localidad sería un gran mal entendido.

Ahora vamos por otro lado. Supongamos que lo que me gustan son las propensiones de caso individual. Ya dijimos que fácilmente puedo relacionar la separabilidad con la causalidad. Ya dijimos que inmediatamente puedo negar que la desigualdad de Bell 2 es la correcta y sé que la desigualdad de Bell 1 es el camino a seguir. Inmediatamente puedo llegar a una conclusión de no-localidad cuando veo que los experimentos confirman la violación de las desigualdades. Conclusión: la no localidad es ineludible.

De esta manera se puede llegar a conclusiones en oposición sobre este tema, por ejemplo [4, 22, 29, 28, 34]. Es una situación que me parece relacionada al famoso concepto de inconmensurabilidad de Kuhn[46]. Pareciera casi imposible comparar ambas aproximaciones pues sus lenguajes y conclusiones están demasiado separadas y tienen pocos puntos de contacto como para establecer comunicación.

¿Cómo podemos afirmar que la comparación no se puede hacer si justo unos párrafos antes hemos comparado cómo distintas aproximaciones a la probabilidad llevan a las diferentes conclusiones? Muy simple, tenemos que agregar otro elemento sobre el que insistimos en la primera parte de este trabajo: nuestro abuso de la probabilidad.

En general, en la mayoría de los textos que hemos venido citando, ninguno de los autores se molesta por especificar aunque sea levemente la teoría de la probabilidad a la que se adscriben o a que usarán en ese trabajo. Muy probablemente porque ni siquiera tengan una.

Si ciñéndonos a entornos de probabilidad diferentes hemos obtenido conclusiones opuestas, imagínese que podemos lograr si no nos sujetamos a ningún entorno de probabilidad y, de manera inconsistente, vamos tomando ideas de uno y de otro conforme se nos presente la oportunidad.

Por eso hemos traído a colación la inconmensurabilidad, el conocimiento de las distintas teorías de la probabilidad es lo que ha permitido que veamos cómo llegar a conclusiones diferentes y en principio podría ayudarnos a establecer puntos de contacto entre distintas visiones del tema. Sin embargo, dicho conocimiento no es parte del adoctrinamiento científico estándar, para muestra basta abrir cualquier libro universitario de probabilidad y ver que difícilmente menciona que hay muchas escuelas de pensamiento sobre el tema. Luego no hay puntos de contacto.

Esto va más allá. Incluso saber de las distintas teorías de la probabilidad no parece ser suficiente. Buena parte de las exégesis de las desigualdades no trabajan sobre un entorno de probabilidad consistente; insisto toman partes de uno y otro conforme el discurso lo necesita. Así lo hace toda la ciencia de hoy. Si les exigiéramos que se ciñeran estrictamente a un entorno de probabilidad probablemente morirían, perderían fuerza o limitaría su alcance. Matar un discurso no es exactamente la mejor forma de encontrar puntos de contacto con él. Así, me parece que al menos tenemos un símil con las distintas escuelas de pensamiento de los electricistas de finales del siglo XVIII y principios del XIX, al tener distintos conceptos de electricidad, justo el ejemplo que Kuhn utiliza para hablar de incomensurabilidad.

Al igual que otros capítulos le invito a pensar en alguna discusión en la literatura de la mecánica cuántica o los teoremas no-go. ¿Cuántas de esas discusiones tienen el único objetivo de atraer la atención, de jalar legitimidad, de verse bien ante el público? Realmente es difícil y tal vez hasta imposible el "verdadero" debate, el que busca entender al otro y ya no digamos encontrar acuerdos sino cuando menos encontrar cuáles son esos puntos de diferencia. Pero probablemente pecamos de frívolos al querer un debate de cuento de hadas donde todo mundo pone de su parte y al final llegamos a una verdad y vivimos felices para siempre. Cualquiera que se haya tomado la molestia de abrir el arte de Cicerón sabe que solo se discute para ganarse el favor, convencer o imponerse frente al otro. Lo demás es demagogia.

Parte III

Prisionero iluso de esta ciencia cotidiana...

Capítulo 11

Un comentario sobre el teorema Kochen-Specker y el teorema PBR

Durante mucho tiempo, creí que la discusión de las desigualdades no tenía mucho sentido pues no encontraba ninguna forma de que se pudiera escapar a las conclusiones de las desigualdades de Bell. Después, esta visión se volvió aún más severa pues al estudiar más profundamente el teorema de Kochen-Specker, sentí que éste era aún más potente y sus conclusiones me parecían aún menos evitables. Luego, aunque siempre he considerado a Einstein el físico más relevante de la historia moderna, no pude sin no aceptar conclusiones como la no-localidad.

La principal razón por la que el teorema Kochen Specker me parece el más fuerte es justo por la ausencia de probabilidades. Justo el uso de la probabilidad me parece que representa un problema para la consistencia de cualquier discurso, desde la mecánica cuántica hasta la teoría de la evolución. Tener un teorema que concluye que las asignaciones de variables no son posibles como se espera en una teoría clásica y sin usar probabilidades me parece la herramienta ideal.

Sin embargo, esta visión cambió un poco cuando me enteré los pocos trabajos que intentan de alguna manera escapar a las conclusiones de los teoremas antes mencionados. Realmente son muy pocos los intentos serios de hacer algo así. Este trabajo justo empezó mencionando uno de ellos. Pitowsky en su tesis doctoral empezó por dar un modelo clásico que reproduce las estadísticas de espín[60] y hubo una discusión que acompaño las fuertes aseveraciones de Pitowsky [50, 52, 59, 58]. No explicaré el método, pero me gustaría resaltar que el trabajo de Pitowsky usa matemáticas poco familiares para el físico común. El uso de inducción transfinita es probablemente el elemento más llamativo. Francamente no conozco otro trabajo en la física que se acerque a matemáticas que, dicho sea de paso, no son familiares ni siquiera al grueso de los matemáticos sino solo a aquellos que trabajan en los fundamentos de las matemáticas. Si este trabajo usa matemáticas tan fuera del área de confort que los físicos tenemos, tampoco me parece tan raro que haya llegado a conclusiones que no se habían visto antes.

Este trabajo tuvo una continuidad marginal principalmente durante la última década del siglo XX y llegó hasta un trabajo también poco conocido, pero del mismo corte llamado *Finite Precision Measurement Nullifies the Kochen-Specker Theorem* [53]. Este trabajo realmente cumple lo que promete. Así como el teorema de Kochen Specker trata sobre las bases en R^3, este trabajo demuestra cómo al limitarse a un espacio de tres dimensiones usando solo números racionales se puede burlar el teorema. Naturalmente, desde el punto de vista empírico no se puede diferenciar entre los reales y los racionales, por lo que ambas representaciones serían equivalentes. Esto también trajo una discusión fuerte a favor y en contra de las aseveraciones de Meyer[7, 9, 13, 41, 42] y muchas otras referencias.

El trabajo de Meyer no tiene errores matemáticos de ningún tipo. Realmente representa una salida genuina al teorema de Kochen Specker. Además una salida que no parece tan artificial como se podría pensar, la finitud de nuestros métodos de medición realmente parece una buena hipótesis de trabajo para motivar el uso de los racionales u otros campos en vez de los reales. Sobre la esta situación, Budroni, Cabello y otros publicaron un artículo muy largo en el 2023 que habla de la contextualidad en todos los aspectos básicos. Sobre la situación con Meyer escriben lo siguiente[12]:

> The KS theorem was developed in the framework of ideal measurements and, as we saw in Sec. IV.B and IV.C, several problems arise when one tries to map those ideal measurements to actual experimental implementations. Some of the first criticisms regarding the physical implications of the KS theorem precisely involved this transition from ideal to actual measurements, in particular the impossibility of arbitrarily precise measurements, and were raised by Meyer (1999), Kent (1999), Clifton and Kent (2000), and Barrett and Kent (2004). These works played a fundamental role in the development of the modern approach to contextuality by stimulating the extension of the KS notion of contextuality from a logical to a probabilistic framework. In fact, they motiva-

> ted the derivation of the KS inequalities, which appeared in those years (Larsson, 2002; Simon et al., 2001). This transition from the logical to the probabilistic perspective in Kochen-Specker's contextuality, and in particular the subsequent theoretical and experimental effort in testing contextuality on physical systems, is the most interesting outcome of this debate.

Pongo esta cita aquí porque me parece que sugiere la aceptación de la crítica de Meyer y otros. Pero también la pongo por una razón mucho más importante: note cómo se acepta que la fundamentación de la contextualidad pasó de un marco "lógico" a uno "probabilístico". Efectivamente, en este mismo trabajo usted puede constatar el desarrollo de una variedad grande de desigualdades y otros objetos matemáticos que, basados en probabilidad, caracterizan la contextualidad. Esto ya viene desde hace tiempo[14] y probablemente desde antes.

Así llegamos a la cuestión importante de esta sección: el teorema de Kochen Specker se ha quedado atrás y la caracterización actual de la contextualidad se hace en términos probabilísticos. La pregunta obligada de este trabajo es ¿Cuál teoría de la probabilidad se usa para todos estos trabajos? La respuesta es la de siempre: ninguna, se usa la misma práctica probabilística que usamos todos los días en la ciencia, a veces consistente, a veces no.

Este comentario se extiende sin muchos cambios al teorema PBR[63] que me parece sigue de cerca en importancia, al igual que los trabajos de contextualidad no establece ningún entorno de probabilidad y sugiere en su discurso algunas veces uno otras otro. No será difícil encontrar distintas interpretaciones del resultado[11, 23, 47, 66, 70] En general, me parece que se lleva bien con las propensiones de caso individual, al igual que el teorema de Bell, pero claramente esa probabilidad le quita sorpresa a su conclusión.

Para terminar quisiera decir que no es la intención de este comentario demeritar los teoremas no-go. Son trabajos realmente sorprendentes e interesantes. Tal vez sí es la intención reafirmar algo que viene desde el principio del trabajo. No se trata de desconfiar de estos teoremas, sino de cualquier discurso que en este u otro mundo haga uso de la probabilidad. Muchas veces este uso no tiene sentido, pero el peligro no es ese. El peligro es que la mayor parte de las veces hay muchas formas de darle sentido y no se especifica cuál es la que se intenta usar.

Este comentario sí es una invitación a desconfiar cada vez que escuche la palabra probabilidad y piense durante un segundo ¿a qué se referirán con la palabra probabilidad? ¿tendrá sentido lo que dicen? ¿este discurso podría incorporar una teoría de la probabilidad bien formulada o moriría en el intento?

Es momento de pasar al final de este trabajo, pues sin importar cuanto podamos argumentar que no hay una teoría de la probabilidad, que nuestro uso de la probabilidad no sigue coherencia, que desconfiemos de todo lo que la probabilidad toca; decirlo la primera vez es sorprendente, la segunda interesante, pero de la tercera en adelante ya es más bien necedad.

Capítulo 12

La Virgen de Guadalupe y la probabilidad en la ciencia

En el siglo XX la probabilidad se convirtió sin lugar a dudas en el instrumento más usado de la ciencia. Esto no es una exageración ni un sentido figurado. Pocas disciplinas no empezaron a hacer uso de los métodos estadísticos. En el caso de la física, son las técnicas de la mecánica estadística las que permitieron una nueva comunicación con la química. La mecánica cuántica se basará fundamentalmente en la probabilidad y se convertirá en la herramienta por antonomasia de la física en el siglo XX a la que se le atribuirán, casi siempre de manera errónea, buena parte de los avances tecnológicos que dieron lugar a lo que conocemos como la era moderna.

De lo anterior discutido parece obvia la conclusión que llegamos. El uso de la probabilidad en el siglo XX no está justificado por el discurso probabilístico de su tiempo. Es un uso que se ha caracterizado por ignorante, alevoso, deliberadamente ambiguo a fin de tener margen de maniobra y toda una serie de características que sin lugar a dudas erizarían la piel de cualquiera de los miembros de la liga de la de-ciencia. La escuela de Copenhagen es un excelente ejemplo de estas estrategias como se puede ver en el joven Feyerabend [26, 27]

En este contexto nos atrevemos a decir que la discusión en mecánica cuántica es secundaria. Al fin y al cabo la imagen pública que sostiene la mecánica cuántica se ha alejado de la realidad conforme pasa el tiempo y se ha vuelto un artículo de novedad. Sobre todo la ciencia ficción y la divulgación de la ciencia han abierto el paso para nuevas generaciones de interpretaciones del gato de Schrodinger que seguramente hacen que todos los días el pobre Schrodinger se revuelque en su tumba.

Entonces, dejemos de lado la mecánica cuántica por un momento sin importarnos si su uso de la probabilidad tiene siquiera sentido, olvidémonos de la mecánica estadística y sus grandes pecados de probabilidad, seamos sinceros: aún si la comunidad de física aceptara dicho uso pecaminoso tendría un impacto despreciable en el discurso científico y mucho menos que despreciable en el discurso público. Esto no es una exageración, para ejemplo pensemos en el antes citado anuncio que la sociedad de estadística americana(sí estadounidense) hizo sobre los valores P, una de las técnicas de pruebas de hipótesis más usadas en el mundo, cuando no la que más. La comunidad científica difícilmente se sacudió por semejante mensaje. Además, el elemento que más llama mi atención y raya en el absurdo es que el grueso de la comunidad científica que usa estas herramientas continuó usándolas sin siquiera inmutarse, como si nada hubiera pasado. Muchos porque francamente ni siquiera se enteraron. Al fin y al cabo por qué alguien que trabaja en, digamos, biología, física experimental o análisis de datos, le debe de importar en lo más mínimo lo que la sociedad de estadística con más poder público tenga que decir. Es mucho más entretenido pensar en historias, que naturalmente no podría repetir en nombre y referencia, que al enterarse de la situación y analizarla a conciencia decían francamente: "sí, me enteré, la leí y tienen razón, pero no tenemos otra forma de hacerlo". Otros ya en pleno ejercicio del iconoclasia: "Yo llevo décadas usándolos y siempre me han funcionado bien, por qué los voy a dejar de usar". Y en general, se pueden contar distintos argumentos que la comunidad científica se inventa para poder continuar, como dicta el progreso, con sus tareas. Al fin detrás del argumento "...no tenemos otra forma de hacerlo..." se implica que hay que hacer algo. Lo que sea, no importa si está mal o no tiene fundamento.

Dicha situación no deja de dar en mi cabeza golpes con los problemas que las apariciones de la Virgen de Guadalupe enfrentaron a mediados del siglo XIX. Nos referimos a la "carta anti-aparicionista" del Sr. García Icabalzeta donde esgrime una larga lista de argumentos en contra de las apariciones marianas en el cerro del Tepeyac que han sido parte fundamental de la identidad mexicana. Para él parece difícil sostener esta tesis en contra de la ideología dominante de la época. Naturalmente son sus palabras mejores que otras para describir la situación. Aquí leamos la despedida de la antedicha acta que es en realidad una carta dirigida al Ilmo. Sr. Arzobispo de México D. Pelagio Antonio de Labastida y Dávalos

> Católico soy; aunque no bueno, Ilmo. Sr., y devoto en cuanto puedo, de la Santísima Virgen; á nadie querría quitar esta devoción: la imagen de Guadalupe será siempre la más antigua, devota y

> respetable de México. Si contra mi intención, por pura ignorancia, se me hubiese escapado alguna palabra o frase mal sonante, desde ahora la doy por no escrita. Por supuesto, que no niego la posibilidad y realidad de los milagros: el que estableció las leyes, bien puede suspenderías ó derogarlas; pero la Omnipotencia Divina no es una cantidad matemática susceptible de aumento o disminución, y nada le añade o le quita un milagro más ó menos. De todo corazón quisiera yo que uno tan honorífico para nuestra patria fuera cierto, pero no lo encuentro así; y si estamos obligados a creer y pregonar los milagros verdaderos, también nos está prohibido divulgar y sostener los falsos. Cuando no se admita que el de la Aparición de Ntra. Sra.de Guadalupe (como se cuenta), es de estos últimos, a lo menos, no podrá negarse que está sujeto a gravísimas objeciones. Si éstas no se destruyen (lo cual hasta ahora no se ha hecho), las apologías producirán efecto contrario. En mi juventud creí, como todos los mexicanos, en la verdad del milagro: no recuerdo de dónde me vinieron las dudas, y para quitármelas acudí a las apologías: éstas convirtieron mis dudas en certeza de la falsedad del hecho. Y no he sido el único. Por eso juzgo que es cosa muy delicada seguir defendiendo la historia. Si he escrito aquí acerca de ella, ha sido por obedecer el precepto repetido de V.S.I. Le ruego, por lo mismo, con todo el encarecimiento que puedo, que este escrito, hijo de la obediencia, no se presente a otros ojos ni pase a otras manos: así me lo ha prometido V.S.I. Me repito de V.S.I. afectísimo amigo y obediente servidor, que su pastoral anillo besa.

Le pido a quién esté leyendo esto que no malinterprete esta cita como una comparación, que, dicen, siempre son malas. Este escrito es una triste tesis doctoral con importancia nula y jamás se acercaría a una situación de verdadero conflicto moral, religioso, político, ético y otras muchas clasificaciones vacías que sí enfrentó un personaje como García Icabalzeta Pero aceptará usted que hay en la carta un comentario particular justo al final al pedir que dicha carta no se haga pública. Parece que él no desea que su estudio salga a la luz. Por un lado bien podría ser miedo de las consecuencias que le hubiera podido traer a él en lo personal. Pero que lo pida en confianza a alguien con tanto poder dentro del clero sugiere más. Pareciera que él no quiere contrariar estas apariciones por la importancia que ellas tienen para los mexicanos. Pareciera que, por encima de que sus estudios señalen la incosistencia de la fuentes, prefiere que esa realidad mexicana continúe y no intentar dañar uno de los pilares en los que se apoya.

¿No es acaso lo que la ciencia hace hoy día? Toda la información que pueden encontrar en este trabajo lleva años publicada y varios libros se han escrito sobre las teorías de la probabilidad. Sin embargo, la ciencia sigue haciendo uso de la probabilidad sin importar que nuestro uso no tenga sentido. Alrededor del mundo se adoctrina en ciencias usando libros de texto que omiten dilemas fundamentales para la probabilidad con una facilidad que me deja perplejo. Incluso buena parte de la comunidad es consciente de dichos problemas y los ignora en su práctica cotidiana como ya hemos mencionado. Parece entonces que al igual que Icabalzeta calla su estudio en favor de conservar un elemento tan fundamental de la realidad mexicana, nosotros privilegiamos la ciencia y su progreso, su continuidad, por encima de esos ideales que la imagen pública de la ciencia usa todos los días en su batalla: la verdad, el pensamiento crítico y otros conceptos vanos.

¿Qué nos dice que continuemos así? Que deseamos que la ciencia continue sin importar si persigue o no algún tipo verdad, qué nos dice, que nos importa más que la ciencia continue sin importar de que se vea comprometida su coherencia. La ciencia de hoy tiene un objetivo: la ciencia, más ciencia, la misma ciencia, que se siga haciendo ciencia. Queremos continuar en ese camino infinito pavimentado con artículo tras articulo, libro de texto tras libro de texto, doctorado tras doctorado, valor promedio tras valor promedio, valores p, intervalos de confianza y demás instrumentos de los que ya no tenemos vínculo con eso que intentaban describir originalmente.

La ley de hoy es publicar. Publicar para poner otro ladrillo en el pavimento de ese camino infinito llamado progreso científico. Ya no sabemos a donde vamos, ya no sabemos cómo llegamos aquí, ni por qué lo estamos construyendo.

La nueva máxima reza “cállate y déjame publicar”.

Parte IV

Apéndices

Apéndice A

Equivalencia de las desigualdades

La forma de las desigualdades Bell-CHSH que podemos encontrar en los trabajos de Fine, Berkovitz y otros no corresponden a lo que normalmente se maneja en la literatura, tamapoco con las de los teoremas originales. Esto se debe a que estos autores trabajan con binarias que se corresponden mejor con la probabilidad. Específicamente, mientras que normalmente pensamos en mediciones de espín con valores -1 o $+1$, estos trabajos cercanos a la probabilidad piensan en los valores 0 y $+1$ respectivamente. Veamos que esta definición de valores nos lleva a las desigualdades que estamos acostumbrados.

Para comenzar hagamos un poco de notación compatible con la de Fine. Si una variable S tiene valores ± 1, entonces definimos $\tilde{S} = \frac{S+1}{2}$ que corresponde con los valores deseados, es decir:

$$\tilde{S}(x) = \begin{cases} 1 & \text{si } x = 1 \\ 0 & \text{si } x = -1 \end{cases} \tag{A.1}$$

$\tilde{S} = 1$ cuando $S = 1$ y $\tilde{S} = 0$ cuando $S = -1$.

Continuamos con la definición de las probabilidades P, que se pueden llamar así pues corresponderán a promedios de funciones características:

$$P(S) = \int \tilde{S}(\lambda)\rho(\lambda)d\lambda, \tag{A.2}$$

$$P(ST) = \int \tilde{S}(\lambda)\tilde{T}(\lambda)\rho(\lambda)d\lambda. \tag{A.3}$$

Estas integrales toman la siguiente forma al cambiar la variable a S:

$$\begin{aligned}
P(S) &= \frac{1}{2}\int S\rho(\lambda)d\lambda + 1 = \frac{1}{2}(\langle S\rangle + 1),\\
P(ST) &= \frac{1}{4}\int (S+1)(T+1)\rho(\lambda)d\lambda = \frac{1}{4}(\langle ST\rangle + \langle S\rangle + \langle T\rangle + 1).
\end{aligned}$$

De lo anterior podemos tomar el término central la desigualdad de Bell en la forma como se encuentra en Fine o en Berkovitz y hacer una sustitución directa:

$$\begin{aligned}
G &= P(XY) + P(X'Y) + P(XY') - P(X'Y') - P(X) - P(Y)\\
&= \frac{1}{4}\Big(\langle XY\rangle + \langle X\rangle + \langle Y\rangle + 1\\
&\qquad \langle X'Y\rangle + \langle X'\rangle + \langle Y\rangle + 1\\
&\qquad \langle XY'\rangle + \langle X\rangle + \langle Y'\rangle + 1\\
&\qquad - \langle X'Y'\rangle - \langle X'\rangle - \langle Y'\rangle - 1\Big)\\
&+ \frac{1}{2}\Big(- \langle X\rangle - 1\\
&\qquad - \langle Y\rangle - 1\Big)\\
&= \frac{1}{4}\Big(\langle XY\rangle + \langle X'Y\rangle + \langle XY'\rangle - \langle X'Y'\rangle - 2\Big)\\
&= \frac{S-2}{4}.
\end{aligned}$$

Finalmente sustituimos en la desigualdad mencionada:

$$\begin{aligned}
-1 \leq& G \leq 0\\
-4 \leq& S - 2 \leq 0\\
-2 \leq& S \leq 2
\end{aligned} \tag{A.4}$$

que es la forma usual de la desigualdad de Bell como buscábamos.

Empecemos por desarrollar un par de desigualdades parciales:

$$\begin{aligned}
&P(A_0 = 1, B_0 = 1, A_{45} = 1)\\
&\quad = P(A_0 = 1, B_0 = 1, A_{45} = 1, B_{45} = 1) + P(A_0 = 1, B_0 = 1, A_{45} = 1, B_{45} = 0)\\
&\quad \leq P(B_0 = 1, B_{45} = 1) + P(A_{45} = 1, B_{45} = 0)\\
&\quad = P(B_0 = 1, B_{45} = 1) + P(A_{45} = 1) - P(A_{45} = 1, B_{45} = 1)
\end{aligned} \tag{A.5}$$

Esta misma expresión se puede usar para otro caso:

$$\begin{aligned}
&P(A_0 = 0, B_0 = 1, A_{45} = 1) \\
&\quad = P(A_0 = 0, B_0 = 1, A_{45} = 1, B_{45} = 1) + P(A_0 = 0, B_0 = 1, A_{45} = 1, B_{45} = 0) \\
&\quad \leq P(A_{45} = 1, B_{45} = 1) + P(B_0 = 1, B_{45} = 0) \\
&\quad = P(A_{45} = 1, B_{45} = 1) + P(B_0 = 1) + P(B_0 = 1, B_{45} = 1)
\end{aligned} \tag{A.6}$$

Por otro lado:

$$\begin{aligned}
&0 \leq P(A_0 = 1, B_0 = 0, A_{45} = 0) \\
&\quad = P(A_0 = 1, B_0 = 0) - P(A_0 = 1, B_0 = 0, A_{45} = 1) \\
&\quad = P(A_0 = 1) - P(A_0 = 1, B_0 = 1) - P(A_0 = 1, A_{45} = 1) + P(A_0 = 1, B_0 = 1, A_{45} = 1)
\end{aligned} \tag{A.7}$$

De la expresión anterior y

$$-P(B_0 = 1, A_{45} = 1) + P(B_0 = 0, A_{45} = 0) = 1 - P(B_0 = 1) - P(A_{45} = 1)$$

tenemos:

$$\begin{aligned}
0 &\leq P(A_0 = 0, B_0 = 0, A_{45} = 0) \\
&= P(B_0 = 0, A_{45} = 0) - P(A_0 = 1, B_0 = 0, A_{45} = 0) \\
&= 1 - P(A_0 = 1) - P(B_0 = 1) - P(A_{45} = 1) + P(A_0 = 1, B_0 = 1) \\
&\qquad + P(A_0 = 1, A_{45} = 1) + P(A_0 = 0, B_0 = 1, A_{45} = 1)
\end{aligned} \tag{A.8}$$

Utilizando las inecuaciones A.5 y A.7 obtenemos:

$$\begin{aligned}
0 \leq\ & P(A_0 = 1) - P(A_0 = 1, B_0 = 1) - P(A_0 = 1, B_{45} = 1) \\
&+ P(B_0 = 1, B_{45} = 1) + P(A_{45} = 1) - P(A_{45} = 1, B_{45} = 1)
\end{aligned} \tag{A.9}$$

De manera análoga usamos A.6 y A.8 para obtener:

$$\begin{aligned}
0 \leq\ & 1 - P(A_0 = 1) - P(A_{45} = 1) + P(A_0 = 1, B_0 = 1) \\
&+ P(A_0 = 1, B_{45} = 1) + P(B_0 = 1, B_{45} = 1) - P(A_{45} = 1, B_{45} = 1)
\end{aligned} \tag{A.10}$$

De las últimas dos se obtiene inmediatamente la desigualdad deseada:

$$\begin{aligned}
-1 \leq\ & P(A_0 = 1, B_0 = 1) + P(A_0 = 1, B_{45} = 1) + P(B_0 = 1, B_{45} = 1) \\
&- P(A_{45} = 1, B_{45} = 1) - P(A_0 = 1) - P(A_{45} = 1) \leq 0
\end{aligned} \tag{A.11}$$

Bibliografía

[1] Full article: The ASA Statement on p-Values: Context, Process, and Purpose. https://www.tandfonline.com/doi/full/10.1080/00031305.2016.1154108.

[2] Quantum logic. In I. Pitowsky, editor, *Quantum Probability — Quantum Logic*, Lecture Notes in Physics, pages 100–137. Springer, Berlin, Heidelberg, 1989.

[3] Soviet Mathematics and Dialectics in the Stalin Era. *Historia Mathematica*, 27(1):54–76, Feb. 2000.

[4] Locality in quantum mechanics: Reply to critics. In J. S. Bell, editor, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*, pages 63–66. Cambridge University Press, Cambridge, 2 edition, 2004.

[5] Probability axioms. *Wikipedia*, Aug. 2023.

[6] C. Abellán, A. Acín, A. Alarcón, O. Alibart, C. K. Andersen, F. Andreoli, A. Beckert, F. A. Beduini, A. Bendersky, M. Bentivegna, P. Bierhorst, D. Burchardt, A. Cabello, J. Cariñe, S. Carrasco, G. Carvacho, D. Cavalcanti, R. Chaves, J. Cortes-Vega, A. Cuevas, A. Delgado, H. de Riedmatten, C. Eichler, P. Farrera, J. Fuenzalida, M. Garcia-Matos, R. Garthoff, S. Gasparinetti, T. Gerrits, F. Ghafari Jouneghani, S. Glancy, E. S. Gómez, P. González, J.-Y. Guan, J. Handsteiner, J. Heinsoo, G. Heinze, A. Hirschmann, O. Jiménez, F. Kaiser, E. Knill, L. T. Knoll, S. Krinner, P. Kurpiers, M. A. Larotonda, J.-Å. Larsson, A. Lenhard, H. Li, M.-H. Li, G. Lima, B. Liu, Y. Liu, I. H. López Grande, T. Lunghi, X. Ma, O. S. Magaña-Loaiza, P. Magnard, A. Magnoni, M. Marti-Prieto, D. Martinez, P. Mataloni, A. Mattar, M. Mazzera, R. P. Mirin, M. W. Mitchell, S. Nam, M. Oppliger, J.-W. Pan, R. B. Patel, G. J. Pryde, D. Rauch, K. Redeker, D. Rieländer, M. Ringbauer,

T. Roberson, W. Rosenfeld, Y. Salathé, L. Santodonato, G. Sauder, T. Scheidl, C. T. Schmiegelow, F. Sciarrino, A. Seri, L. K. Shalm, S.-C. Shi, S. Slussarenko, M. J. Stevens, S. Tanzilli, F. Toledo, J. Tura, R. Ursin, P. Vergyris, V. B. Verma, T. Walter, A. Wallraff, Z. Wang, H. Weinfurter, M. M. Weston, A. G. White, C. Wu, G. B. Xavier, L. You, X. Yuan, A. Zeilinger, Q. Zhang, W. Zhang, J. Zhong, and The BIG Bell Test Collaboration. Challenging local realism with human choices. *Nature*, 557(7704):212–216, May 2018.

[7] D. M. Appleby. Existential contextuality and the models of Meyer, Kent, and Clifton. *Physical Review A*, 65(2):022105, Jan. 2002.

[8] L. E. Ballentine. Einstein's Interpretation of Quantum Mechanics. *American Journal of Physics*, 40(12):1763–1771, Dec. 1972.

[9] J. Barrett and A. Kent. Non-contextuality, finite precision measurement and the Kochen–Specker theorem. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 35(2):151–176, June 2004.

[10] J. S. Bell. The theory of local beables. In *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*, pages 52–62. Cambridge University Press, Cambridge, 2 edition, 2004.

[11] Y. Ben-Menahem. The PBR theorem: Whose side is it on? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 57:80–88, Feb. 2017.

[12] C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, and J.-Å. Larsson. Kochen-Specker contextuality. *Reviews of Modern Physics*, 94(4):045007, Dec. 2022.

[13] A. Cabello. Finite-precision measurement does not nullify the Kochen-Specker theorem. *Physical Review A*, 65(5):052101, Apr. 2002.

[14] A. Cabello. Experimentally Testable State-Independent Quantum Contextuality. *Physical Review Letters*, 101(21):210401, Nov. 2008.

[15] R. Carnap. *Logical Foundations of Probability, 2Nd Edition.* Chicago Univeresity Press, 1962.

[16] L. Chaumont, L. Mazliak, and M. Yor. Some aspects of the probabilistic work. In É. Charpentier, A. Lesne, and N. K. Nikolski, editors,

Kolmogorov's Heritage in Mathematics, pages 41–66. Springer, Berlin, Heidelberg, 2007.

[17] K. L. Chung and K. Zhong. *A Course in Probability Theory.* Academic Press, 2001.

[18] A. I. Dale and P.-S. Laplace. *Philosophical Essay on Probabilities*, volume 13 of *Sources in the History of Mathematics and Physical Sciences.* Springer, New York, NY, 1995.

[19] L. Daston. *Classical Probability in the Enlightenment.* Princeton University Press, 1988.

[20] B. de Finetti. *Theory of Probability: A Critical Introductory Treatment.* Wiley, 1979.

[21] B. de Finetti. Probabilism: A Critical Essay on the Theory of Probability and on the Value of Science. *Erkenntnis (1975-)*, 31(2/3):169–223, 1989.

[22] L. de la Pena, A. M. Cetto, and T. A. Brody. On hidden-variable theories and Bell's inequality. *Lettere al Nuovo Cimento (1971-1985)*, 5(2):177–181, Sept. 1972.

[23] A. Drezet. The PBR theorem seen from the eyes of a Bohmian. *arXiv:1409.3478 [quant-ph]*, Sept. 2014.

[24] A. Einstein, M. Born, and H. Born. *The Born-Einstein Letters: Correspondence Between Albert Einstein and Max and Hedwig Born from 1916-1955, with Commentaries by Max Born.* Macmillan, 1971.

[25] B. Ellis. *Basic Concepts of Measurement.* Cambridge University Press, Jan. 1966.

[26] P. K. Feyerabend. On a Recent Critique of Complementarity: Part I. *Philosophy of Science*, 35(4):309–331, Dec. 1968.

[27] P. K. Feyerabend. On a Recent Critique of Complementarity: Part II. *Philosophy of Science*, 36(1):82–105, Mar. 1969.

[28] A. Fine. Fine Responds. *Physical Review Letters*, 49(3):243–243, July 1982.

[29] A. Fine. Hidden Variables, Joint Probability, and the Bell Inequalities. *Physical Review Letters*, 48(5):291–295, Feb. 1982.

[30] R. A. Fisher. Statistical Methods for Research Workers. In S. Kotz and N. L. Johnson, editors, *Breakthroughs in Statistics: Methodology and Distribution*, Springer Series in Statistics, pages 66–70. Springer, New York, NY, 1992.

[31] J. I. B. Fuentes and J. G. Aranzabal. Una propuesta para la enseñanza de la probabilidad en la universidad basada en la investigación didáctica. *Educación Matemática*, 21(3):127–162, 2009.

[32] M. C. Galavotti. The notion of subjective probability in the work of Ramsey and de Finetti. *Theoria*, 57(3):239–259, 1991.

[33] M. C. Galavotti. Operationism, Probability and Quantum Mechanics. In B. C. Van Fraassen, editor, *Topics in the Foundation of Statistics*, pages 99–118. Springer Netherlands, Dordrecht, 1997.

[34] A. Garg and N. D. Mermin. Correlation Inequalities and Hidden Variables. *Physical Review Letters*, 49(17):1220–1223, Oct. 1982.

[35] B. Gerrard. Ramsey and Keynes revisited. *Cambridge Journal of Economics*, 47(1):195–213, Jan. 2023.

[36] D. Gillies. *Philosophical Theories of Probability*. Routhledge, 2000.

[37] N. Gisin. Propensities in a Non-Deterministic Physics. *Synthese*, 89(2):287–297, 1991.

[38] K. Gödel. *Kurt Gödel: Collected Works: Volume I: Publications 1929-1936*. OUP USA, May 1986.

[39] A. Gut. *Probability: A Graduate Course*. Springer Texts in Statistics. Springer-Verlag, New York, 2005.

[40] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, Oct. 2015.

[41] R. Hermens. Conway–Kochen and the Finite Precision Loophole. *Foundations of Physics*, 44(10):1038–1048, Oct. 2014.

[42] A. Kent. Noncontextual Hidden Variables and Physical Measurements. *Physical Review Letters*, 83(19):3755–3757, Nov. 1999.

[43] J. M. Keynes. *A Treatise on Probability: By John Maynard Keynes.* Macmillan, 1921.

[44] S. C. Kleene. *Mathematical Logic.* Courier Corporation, Jan. 2002.

[45] A. N. Kolmogorov and A. T. Bharucha-Reid. *Foundations of the Theory of Probability: Second English Edition.* Chelsea Publishing Company, Apr. 2018.

[46] T. S. Kuhn. *La estructura de las revoluciones científicas.* Fondo de Cultura Economica, Jan. 2019.

[47] M. S. Leifer. Is the quantum state real? An extended review of ψ-ontology theorems. *Quanta*, 3(1):67, Nov. 2014.

[48] C. López, P. Gómez, C. López, and P. Gómez. Probabilidad en diferentes países del mundo: enseñanza de la probabilidad en educación primaria. *Educación matemática*, 34(3):42–64, 2022.

[49] Y. LU and I. BELITSKAYA-LEVY. The debate about p-values. *Shanghai Archives of Psychiatry*, 27(6):381–385, Dec. 2015.

[50] A. L. Macdonald. Comment on "Resolution of the Einstein-Podolsky-Rosen and Bell Paradoxes". *Physical Review Letters*, 49(16):1215–1215, Oct. 1982.

[51] P. S. marquis de Laplace. *A Philosophical Essay on Probabilities.* Dover, 1951.

[52] N. D. Mermin. Comment on "Resolution of the Einstein-Podolsky-Rosen and Bell Paradoxes". *Physical Review Letters*, 49(16):1214–1214, Oct. 1982.

[53] D. A. Meyer. Finite Precision Measurement Nullifies the Kochen-Specker Theorem. *Physical Review Letters*, 83(19):3751–3754, Nov. 1999.

[54] D. Miller. *Critical Rationalism: A Restatement and Defence.* Open Court, Dec. 2015.

[55] R. V. Mises. *Probability, Statistics, and Truth.* Courier Corporation, Jan. 1981.

[56] D. R. Murdoch. *Niels Bohr's Philosophy of Physics.* Cambridge University Press, Cambridge, 1987.

[57] A. Pais. George Uhlenbeck and the Discovery of Electron Spin. *Physics Today*, 42(12):34–40, Dec. 1989.

[58] I. Pitowsky. Pitowsky Responds. *Physical Review Letters*, 49(16):1216–1216, Oct. 1982.

[59] I. Pitowsky. Resolution of the Einstein-Podolsky-Rosen and Bell Paradoxes. *Physical Review Letters*, 48(19):1299–1302, May 1982.

[60] I. Pitowsky. Deterministic model of spin and statistics. *Physical Review D*, 27(10):2316–2326, May 1983.

[61] I. PITOWSKY. Quantum probability, quantum logic. *Quantum probability, quantum logic*, 321:IX, 1–209 [219 p.], 1989.

[62] K. R. Popper. The propensity interpretation of probability. *The British Journal for the Philosophy of Science*, 10(37):25–42, May 1959.

[63] M. F. Pusey, J. Barrett, and T. Rudolph. On the reality of the quantum state. *Nature Physics*, 8(6):475–478, June 2012.

[64] F. P. Ramsey. Mr Keynes on Probability. *The British Journal for the Philosophy of Science*, 40(2):219–222, 1989.

[65] F. P. Ramsey. *F. P. Ramsey: Philosophical Papers*. Cambridge University Press, July 1990.

[66] A. Rizzi. Does the PBR Theorem Rule out a Statistical Understanding of QM? *Foundations of Physics*, 48(12):1770–1793, Dec. 2018.

[67] S. Rogosin and F. Mainardi. A.Ya. Khintchine's Work in Probability Theory. *Notices of the International Congress of Chinese Mathematicians*, 5(2):60–75, 2017.

[68] S. M. Ross. *A First Course in Probability*. Pearson, 2018.

[69] J. J. Sakurai and J. Napolitano. *Modern Quantum Mechanics*. Cambridge University Press, Sept. 2017.

[70] M. Schlosshauer and A. Fine. Implications of the Pusey-Barrett-Rudolph Quantum No-Go Theorem. *Physical Review Letters*, 108(26):260404, June 2012.

[71] E. Seneta. Mathematics, religion, and Marxism in the Soviet Union in the 1930s. *Historia Mathematica*, 31(3):337, 2004.

[72] A. F. G. Solis-Labastida, M. Gastelum, and J. G. Hirsch. The Violation of Bell-CHSH Inequalities Leads to Different Conclusions Depending on the Description Used. *Entropy*, 23(7):872, July 2021.

[73] L. A. Steen and J. A. Seebach. *Counterexamples in Topology*. Courier Corporation, Apr. 2013.

[74] S.-i. Tomonaga and S. Tomonaga. *The Story of Spin*. University of Chicago Press, 1997.

[75] E. G. Torres. SIGNIFICADOS DE LA PROBABILIDAD EN LIBROS DE TEXTO PARA EDUCACIÓN PRIMARIA EN ANDALUCÍA.

[76] B. C. Van Fraassen. The Charybdis of Realism: Epistemological Implications of Bell's Inequality. *Synthese*, 52(1):25–38, 1982.

[77] R. Vargas and L. Fernando. La educación estadística en el nivel universitario: Retos y oportunidades. *Revista Digital de Investigación en Docencia Universitaria*, 13(2):67–82, July 2019.

[78] V. G. Vovk and G. R. Shafer. Kolmogorov's Contributions to the Foundations of Probability. *Problems of Information Transmission*, 39(1):21–31, Jan. 2003.

[79] F. Yates. The Influence of Statistical Methods for Research Workers on the Development of the Science of Statistics. *Journal of the American Statistical Association*, 46(253):19–34, Mar. 1951.