

La probabilidad de la mecánica cuántica: una introducción en noventa minutos

Stephen Bruce Sontz

sontz@cimat.mx

Centro de Investigación en Matemáticas, A.C.
(CIMAT)
Guanajuato, Gto., México

Resumen

Einstein decía, en referencia a la mecánica cuántica: “The old one does not play dice” - y es cierto, porque la probabilidad que se usa en la mecánica cuántica no es la probabilidad clásica que se usa en los juegos (como los dados). Este es el primer ejemplo, históricamente hablando, de una teoría de probabilidad de las llamadas no clásicas o no conmutativas. En este artículo presentaremos los conceptos de esta probabilidad cuántica a un nivel introductorio, dirigido a estudiantes de licenciatura con conocimiento del álgebra lineal. Usaremos la teoría de probabilidad en un espacio finito como analogía para el caso de dimensión finita de la probabilidad cuántica. Además, ofrecemos un apéndice breve sobre un tema relacionado (qubits) para que el lector vea que las ideas presentadas aquí tienen otras aplicaciones. No se necesita conocimiento alguno de la mecánica cuántica para entender este artículo.

1. Introducción

¿Por qué un artículo que se dice introductorio necesita una introducción? Porque en mi experiencia mucha gente no entiende el papel que un artículo introductorio debe desempeñar. Mi única meta en este artículo

es despertar el interés del lector para que estudie los temas aquí expuestos o quizá algunos temas relacionados a este. Si el lector piensa que hay pocos detalles aquí, espero que siga leyendo otros artículos y textos. Si piensa que hay demasiados detalles, espero que se de cuenta que aquí faltan muchas cosas interesantes. En todo caso, la sección más importante de este artículo es la bibliografía. Por lo tanto, no tienen mucha importancia cosas secundarias como el origen de este artículo en una conferencia, el hecho que su autor no tiene un dominio perfecto del idioma o la exclusión de algunos temas de mucho interés para los expertos en el área. Ojalá que mi meta sea también la meta del lector.

Otra cosa: he aquí *una* introducción. Hay otras, si el lector prefiere otro punto de vista, quiere ver más detalles o tal vez desea leer algo escrito en inglés para aprender la terminología internacional. Véase [1], [5], [9] y [11].

Hay otro asunto que se debe tomar en cuenta, dado que la probabilidad cuántica tiene su origen histórico en la mecánica cuántica. Este artículo no es, de ninguna manera, una introducción a la mecánica cuántica, cosa que implicaría un documento de tal vez cien páginas (o más). Al contrario, no se requiere ningún conocimiento de mecánica cuántica para la lectura y comprensión de este artículo pues la teoría matemática de la probabilidad cuántica es casi completamente auto-suficiente: lo único que faltaría es la motivación física (que viene de la mecánica cuántica). Por lo tanto, los ejemplos presentados en este artículo se estudian desde el punto de vista matemático y no desde el punto de vista de la física.

Una nota para terminar esta introducción: La necesidad de la teoría de la probabilidad cuántica proviene de la física; más concretamente, de la física experimental. No hay, hasta la fecha, una justificación *matemática* para su estudio (de hecho es posible que esta justificación no exista).

2. Probabilidad Clásica - Caso Finito

Antes de empezar le hacemos notar al lector que, dado que este artículo está basado en una conferencia de noventa minutos, no haremos todas las demostraciones. Cada afirmación sin demostración o con demostración incompleta debe considerarse como un ejercicio para el lector.

Un caso especial de la probabilidad clásica consiste de:

1. un conjunto Ω que es *finito* y *no vacío*,
2. *todos* los subconjuntos $E \subset \Omega$, llamados *eventos*,

3. una *función de probabilidad* P tal que $0 \leq P(E) \leq 1$ para cada evento $E \subset \Omega$.

Se dice que $P(E)$ es la *probabilidad del evento* E si se cumplen las siguientes propiedades:

1. $P(\emptyset) = 0$, donde \emptyset denota al conjunto vacío,
2. $P(\Omega) = 1$,
3. $P(E_1 \cup E_2) = P(E_1) + P(E_2)$, si E_1 y E_2 son *eventos disjuntos*, es decir, si $E_1 \cap E_2 = \emptyset$.

Estos son los principios de la teoría de probabilidad clásica desarrollada por Fermat y Pascal, entre otros, en el siglo XVII para estudiar juegos como las cartas. Hoy en día, la probabilidad clásica tiene muchas aplicaciones en biología, ingeniería, física y hasta en las mismas matemáticas.

En este artículo no hablaremos de la teoría de *probabilidad clásica* de Kolmogorov (1933, véase [10]) para el caso cuando Ω es *infinito*, dado que tiene un nivel más avanzado y muchos detalles técnicos.

Una observación clave: existen experimentos con la propiedad de que cuando se llevan a cabo bajo condiciones iniciales iguales (o esencialmente iguales) no dan siempre el mismo resultado; es decir, hay un conjunto de dos (o más) resultados del experimento (o mediciones) posibles. Es para estos casos que surge la necesidad de una teoría que no sea *determinista* (una teoría es determinista si predice exactamente un sólo resultado para cada experimento u observación). Una teoría que describe situaciones donde hay más de un resultado posible se llama una *teoría de probabilidad*. Algunos ejemplos de estas teorías son la probabilidad clásica de Kolmogorov, la probabilidad cuántica y la probabilidad libre (véase [20]).

Por todo esto, en una teoría de probabilidad, las estructuras matemáticas que corresponden a las cantidades *medidas* u *observadas* en experimentos, y expresadas como números reales, son importantes. En la probabilidad clásica, estas estructuras se llaman *variables aleatorias*. Por definición, una variable aleatoria es una función X con valores reales:

$$X : \Omega \rightarrow \mathbb{R}.$$

La idea detrás de esta definición es que los números reales en el rango (o imagen) de X , definido como

$$\text{Ran}(X) := \{\lambda \in \mathbb{R} \mid \exists \omega \in \Omega, \lambda = X(\omega)\},$$

son todos los valores medibles posibles de la cantidad experimental correspondiente a X .

Se define el *valor esperado* (o más bien el *valor promedio*) de X con respecto a P por

$$\langle X \rangle := \sum_{\omega \in \Omega} P(\omega)X(\omega) \quad \text{donde} \quad P(\omega) \equiv P(\{\omega\}).$$

Hacemos notar que el conjunto

$$\{ X \mid X : \Omega \rightarrow \mathbb{R} \}$$

de todas las variables aleatorias es un espacio vectorial sobre \mathbb{R} de dimensión finita $n = \text{card}(\Omega) \geq 1$. Este espacio vectorial también tiene un producto conmutativo dado por la multiplicación usual de funciones:

$$XY(\omega) = (X \cdot Y)(\omega) := X(\omega)Y(\omega)$$

para cada $\omega \in \Omega$ donde X, Y son variables aleatorias.

3. Mecánica Cuántica en Dimensión Finita

La *Mecánica Cuántica* de Pauli y Wigner (~ 1930) empieza con el espacio vectorial

$$\mathbb{C}^n = \{ z = (z_1, z_2, \dots, z_n) \mid z_j \in \mathbb{C} \text{ para } 1 \leq j \leq n \}$$

de *dimensión finita* $n \geq 1$ sobre los números complejos \mathbb{C} , con el *producto interior*

$$\langle z, w \rangle := \sum_{j=1}^n z_j^* w_j.$$

Aquí $z = (z_1, \dots, z_n) \in \mathbb{C}^n$, $w = (w_1, \dots, w_n) \in \mathbb{C}^n$, y β^* es el *complejo conjugado* de $\beta \in \mathbb{C}$. El espacio \mathbb{C}^n tiene una *norma* $\| \cdot \|$ dada por

$$\|z\|^2 = \langle z, z \rangle = \sum_{j=1}^n z_j^* z_j = \sum_{j=1}^n |z_j|^2 \quad \text{para cada } z \in \mathbb{C}^n.$$

No presentaremos en este artículo la teoría de la *mecánica cuántica* de Heisenberg y Schrödinger (1925-26) para el caso de dimensión *infinita*, dado que tiene muchos detalles técnicos. Al igual que la probabilidad clásica general de Kolmogorov, esta teoría es importante pero

no tenemos que estudiarla para alcanzar nuestras metas. Al estudiante interesado en aprender en el caso general la probabilidad clásica, la mecánica cuántica y la probabilidad cuántica (entre otros temas) le recomendamos ampliamente la *lectura* de estos temas en otros artículos y libros. Véase la bibliografía para empezar.

Para definir las estructuras matemáticas en la mecánica cuántica que corresponden a cantidades medidas en un experimento introducimos el conjunto de *matrices* complejas $n \times n$ (con n entero, $n \geq 1$)

$$\text{MAT}(n; \mathbb{C}) := \{A = (A_{jk}), \text{ matriz } n \times n, A_{jk} \in \mathbb{C}, 1 \leq j, k \leq n\}.$$

Cada matriz A define un mapeo lineal $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ de la manera usual:

$$(Az)_j = \sum_{k=1}^n A_{jk} z_k.$$

Recíprocamente cada mapeo lineal de \mathbb{C}^n en \mathbb{C}^n proviene de una matriz única en $\text{MAT}(n; \mathbb{C})$. (Para nosotros *mapeo*, *transformación* y *función* son sinónimos. Usamos un abuso común de notación: el símbolo A denota a la matriz y al mapeo lineal correspondiente.)

Resulta ser que $\text{MAT}(n; \mathbb{C})$ es un espacio vectorial de dimensión n^2 sobre \mathbb{C} . Este espacio cuenta con el producto usual de matrices, que es *no conmutativo* si $n \geq 2$. Por ser un espacio vectorial con un producto compatible con las estructuras vectoriales (suma de vectores y multiplicación escalar), $\text{MAT}(n; \mathbb{C})$ es un *álgebra* sobre \mathbb{C} .

Toda matriz A tiene una *matriz adjunta* A^* donde se define A^* como $(A^*)_{jk} := (A_{kj})^*$ para $1 \leq j, k \leq n$; es decir, A^* es la matriz transpuesta conjugada de A . Resulta ser que A^* es la única matriz que satisface

$$\langle A^* z, w \rangle = \langle z, Aw \rangle$$

para todos los vectores $z, w \in \mathbb{C}^n$.

Si $A = A^*$, se dice que A es *autoadjunta* (o *Hermitiana*). Denotamos las matrices autoadjuntas por

$$\text{HERM}(n) := \{A \in \text{MAT}(n; \mathbb{C}) \mid A = A^*\}.$$

Las matrices autoadjuntas corresponden a muchas de las cantidades medidas en experimentos con sistemas cuánticos. Vamos a ver más adelante con todo detalle esta correspondencia, pero por lo pronto cabe subrayar que es por eso que las matrices autoadjuntas son importantes en física cuántica.

Ejercicio: $\text{HERM}(n)$ es un espacio vectorial sobre los números reales \mathbb{R} y *no* lo es sobre los números complejos \mathbb{C} . El espacio $\text{HERM}(n)$ no es una subálgebra de $\text{MAT}(n; \mathbb{C})$ si $n \geq 2$. Además $\dim_{\mathbb{R}} \text{HERM}(n) = n^2$.

Ejemplo: $\text{HERM}(2)$ tiene dimensión (sobre los reales) $2^2 = 4$ y una base está dada por las tres *matrices de Pauli*

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

y la identidad, $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. (Aquí $i = \sqrt{-1}$.) Algunas propiedades de estas matrices son:

$$\sigma_1\sigma_2 = i\sigma_3 = -\sigma_2\sigma_1, \quad \sigma_2\sigma_3 = i\sigma_1 = -\sigma_3\sigma_2, \quad \sigma_3\sigma_1 = i\sigma_2 = -\sigma_1\sigma_3.$$

4. Teoría Espectral

En esta sección usaremos un teorema fundamental del álgebra lineal: cada matriz autoadjunta tiene una *diagonalización*. Una referencia para esta sección es el texto de Halmos [6].

Sea A una matriz en $\text{MAT}(n; \mathbb{C})$. El *polinomio característico* de A se define como

$$p_A(\lambda) := \det(\lambda I - A),$$

donde $I \in \text{MAT}(n; \mathbb{C})$ es la matriz identidad, \det es el determinante de una matriz $n \times n$ y λ es una variable compleja. El polinomio p_A tiene coeficientes complejos y es de grado $n \geq 1$.

Por el Teorema Fundamental del Álgebra, podemos escribir el conjunto de las raíces complejas *distintas* de $p_A(\lambda)$ como

$$\text{SPEC}(A) := \{\lambda_1, \lambda_2, \dots, \lambda_k\} \subset \mathbb{C}$$

con $1 \leq k \leq n$. Decimos que $\text{SPEC}(A)$ es el *espectro* de A y que $\lambda_1, \lambda_2, \dots, \lambda_k$ son los *eigenvalores* de A . (La palabra alemana *eigen* se suele usar hoy en día y es parte del vocabulario internacional matemático. A veces se usan traducciones como *propio* o *característico*.)

Existen diferentes caracterizaciones de los eigenvalores; mostramos algunas de ellas. Las siguientes afirmaciones son equivalentes:

- $\beta \in \mathbb{C}$ es un eigenvalor de A .
- $p_A(\beta) = 0$.
- $\det(\beta I - A) = 0$.

- La matriz $\beta I - A$ no es invertible.
- El *subespacio* $\ker(\beta I - A)$ no es cero.
(Aquí $\ker(B) := \{z \in \mathbb{C}^n \mid Bz = 0\}$ para $B : \mathbb{C}^n \rightarrow \mathbb{C}^n$ lineal.)
- Existe $z \in \mathbb{C}^n$ con $z \neq 0$ y $Az = \beta z$.
(En tal caso z se llama un *eigenvector de A asociado a β* .)

Quizás la última afirmación sea la propiedad más familiar para el lector. Sin embargo, usaremos la penúltima. Explícitamente, definimos los subespacios

$$V_j := \ker(\lambda_j I - A) \neq 0 \quad \text{para } 1 \leq j \leq k.$$

Entonces $z \in V_j$ si y sólo si $(\lambda_j I - A)z = 0$ si y sólo si $Az = \lambda_j z$; es decir, la acción de A en el subespacio V_j es igual a la acción de multiplicación por λ_j . Se dice que A *tiene el valor λ_j* en el subespacio V_j .

Si además A es autoadjunta, entonces tenemos:

- Todos los eigenvalores de A son reales.
- Los subespacios V_j son ortogonales; es decir,

$$\langle z_i, z_j \rangle = 0 \quad \text{si } z_i \in V_i \text{ y } z_j \in V_j \text{ para } i \neq j.$$

- $\mathbb{C}^n = V_1 \oplus V_2 \oplus \cdots \oplus V_k$. (*Suma Directa*)

Dejamos la tercera afirmación como ejercicio. Para probar las dos primeras afirmaciones, tomamos $z_i \in V_i$ y $z_j \in V_j$. Entonces para $1 \leq i, j \leq k$ tenemos

$$\lambda_j \langle z_i, z_j \rangle = \langle z_i, \lambda_j z_j \rangle = \langle z_i, Az_j \rangle = \langle Az_i, z_j \rangle = \langle \lambda_i z_i, z_j \rangle = \lambda_i^* \langle z_i, z_j \rangle.$$

Tomando $i = j$, $z_i \neq 0$ y usando $\langle z_i, z_i \rangle \neq 0$, concluimos que $\lambda_i = \lambda_i^*$, es decir λ_i es real. Luego tomando $i \neq j$, tenemos que $(\lambda_i - \lambda_j) \langle z_i, z_j \rangle = 0$, que implica $\langle z_i, z_j \rangle = 0$ porque $\lambda_i \neq \lambda_j$.

Al formar la unión de una base ortonormal de V_1 , con una base ortonormal de V_2 , y así sucesivamente hasta llegar a V_k , obtenemos una base ortonormal de \mathbb{C}^n con la propiedad que el mapeo lineal

$$A : \mathbb{C}^n \rightarrow \mathbb{C}^n$$

tiene una matriz diagonal en esta base ortonormal nueva. A esto se le llama la *diagonalización* de la matriz autoadjunta A . En la diagonal λ_1

aparece $\dim V_1$ veces, λ_2 aparece $\dim V_2$ veces, \dots , λ_k aparece $\dim V_k$ veces. Nótese que $\dim V_j \geq 1$ para todo j .

Vayamos un paso más adelante. En lugar de usar subespacios de \mathbb{C}^n vamos a usar proyecciones ortogonales. Sea $W \subset \mathbb{C}^n$ un subespacio de \mathbb{C}^n . Resulta que podemos *descomponer* \mathbb{C}^n como una *suma directa*,

$$\mathbb{C}^n = W \oplus W^\perp,$$

donde el subespacio

$$W^\perp := \{v \in \mathbb{C}^n \mid \langle v, w \rangle = 0 \text{ para todo } w \in W\}$$

se llama el *complemento ortogonal* de W en \mathbb{C}^n .

Sea $z \in \mathbb{C}^n$. Podemos escribir de manera única a z como $z = w + v$ con $w \in W$ y $v \in W^\perp$. Al subespacio W asociamos el mapeo lineal $E_W : \mathbb{C}^n \rightarrow \mathbb{C}^n$ definido por

$$E_W z := w,$$

para cada $z \in \mathbb{C}^n$.

Resulta ser que

$$E_W = E_W^* = E_W^2 \quad \text{y} \quad \text{Ran}(E_W) = W.$$

(Aquí el rango de $B : \mathbb{C}^n \rightarrow \mathbb{C}^n$ es $\text{Ran}(B) := \{u \in \mathbb{C}^n \mid \exists z \in \mathbb{C}^n, u = Bz\}$.) Si $E : \mathbb{C}^n \rightarrow \mathbb{C}^n$ es un mapeo lineal que satisface $E = E^* = E^2$, se dice que E es una *proyección ortogonal*. Al mapeo E_W se le llama la *proyección ortogonal sobre W* .

Recíprocamente, para cada proyección ortogonal E tenemos que $E = E_W$ para un subespacio único, a saber, $W = \text{Ran}(E)$.

A fin de cuentas, hemos obtenido una correspondencia uno a uno y sobre (es decir, una *biyección*) entre el conjunto de todos los subespacios W de \mathbb{C}^n y el conjunto de todas las proyecciones ortogonales E en $\text{MAT}(n; \mathbb{C})$. La correspondencia manda el subespacio W a la proyección ortogonal E_W mientras que la correspondencia inversa manda la proyección ortogonal E al subespacio $\text{Ran}(E)$.

Ahora podemos enunciar, al fin, el teorema de diagonalización. Este teorema usa el concepto de proyección ortogonal para expresar los resultados que ya hemos visto anteriormente.

Teorema: Si $A \in \text{MAT}(n, \mathbb{C})$ es una matriz autoadjunta, entonces existe un entero k con $1 \leq k \leq n$ y existen proyecciones ortogonales $E_1 \neq 0, \dots, E_k \neq 0$ en $\text{MAT}(n, \mathbb{C})$ y números reales *distintos* $\lambda_1, \dots, \lambda_k$

tales que

$$\begin{aligned} E_i E_j &= 0 \quad \text{si } i \neq j, \\ I &= E_1 + E_2 + \cdots + E_k, \\ A &= \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_k E_k. \end{aligned}$$

Además, el entero k , las proyecciones ortogonales y los números reales (con las propiedades indicadas) son únicos. ■

Este resultado fundamental del álgebra lineal se llama *Teorema Espectral* para matrices autoadjuntas. Recíprocamente, si tenemos un entero $k \geq 1$, proyecciones ortogonales y números reales con las propiedades indicadas, entonces podemos definir

$$A := \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_k E_k.$$

La matriz A es autoadjunta.

A la representación de A dada por el Teorema Espectral se le conoce como su *resolución espectral* y es una *forma canónica* de A .

Es importante notar que hay una biyección: $\lambda_j \longleftrightarrow E_j$. Es decir, para cada eigenvalor λ_j (donde $j = 1, \dots, k$) hay una única proyección ortogonal E_j correspondiente y viceversa.

5. Más del Mundo Clásico

Vamos a considerar nuevamente la teoría de probabilidad clásica, cuando Ω es finito y no vacío con $n = \text{card}(\Omega) \geq 1$, pero ahora sin considerar la función de probabilidad P .

Para cada evento $\Lambda \subset \Omega$ definimos su *función característica* χ_Λ por

$$\chi_\Lambda(\omega) := \begin{cases} 1, & \text{si } \omega \in \Lambda, \\ 0, & \text{si } \omega \in \Omega, \omega \notin \Lambda. \end{cases}$$

Se sigue que $\chi_\Lambda = \chi_\Lambda^* = \chi_\Lambda^2$ y además $\Lambda = \chi_\Lambda^{-1}(1)$.

Además, resulta que cada función $\chi : \Omega \rightarrow \mathbb{C}$ que satisface $\chi = \chi^* = \chi^2$ es la función característica de un evento único Λ , a saber, $\chi = \chi_\Lambda$ donde $\Lambda = \chi^{-1}(1)$.

A fin de cuentas tenemos una correspondencia uno a uno y sobre (o sea, una biyección) entre el conjunto de eventos Λ en Ω y el conjunto de las funciones características χ_Λ con dominio Ω .

Ahora supongamos que $A : \Omega \rightarrow \mathbb{R}$ es una función arbitraria (variable aleatoria). Como Ω es un conjunto finito, su rango satisface

$$\text{Ran}(A) = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$$

para un entero k tal que $1 \leq k \leq n$ y para $\lambda_1, \lambda_2, \dots, \lambda_k$ números reales distintos.

Si definimos el espectro de la función A , denotado por $\text{SPEC}(A)$, como el conjunto de números reales λ tal que la función $\lambda - A$ no tiene inversa multiplicativa, entonces $\text{SPEC}(A) = \text{Ran}(A)$.

Para cada “eigenvalor” $\lambda_j \in \text{SPEC}(A)$ definimos el evento

$$\Lambda_j := \ker(\lambda_j - A) = \{\omega \in \Omega \mid A(\omega) = \lambda_j\} = A^{-1}(\lambda_j) \neq \emptyset.$$

Se dice que Λ_j es el evento donde A toma el valor λ_j .

Los eventos Λ_i son no vacíos y son *eventos disjuntos*. Además se tiene

$$\Omega = \bigcup_{j=1}^k \Lambda_j.$$

Definamos $E_j := \chi_{\Lambda_j}$ para cada $j = 1, 2, \dots, k$. De nuevo tenemos una biyección importante: $\lambda_j \longleftrightarrow E_j$. Además,

$$\begin{aligned} E_i E_j &= 0 \quad \text{si } i \neq j, \\ 1 &= E_1 + E_2 + \dots + E_k, \\ A &= \lambda_1 E_1 + \lambda_2 E_2 + \dots + \lambda_k E_k. \end{aligned}$$

A esta descomposición se le llama la *forma canónica* de la función simple A y la demostración es un lema básico en un curso introductorio de la teoría de la medida. Hay que recordar que, en teoría de la medida, una *función simple* es una función (medible) cuyo rango tiene un número finito de valores.

Por cierto, este lema no dice nada sobre medidas, de la misma manera que el teorema espectral para matrices autoadjuntas no dice nada sobre medidas de probabilidad cuántica.

Nótese que el caso conmutativo de funciones tiene la misma estructura algebraica que el caso no conmutativo de matrices. Por esto, a veces se dice que “*la cuantización es usar operadores en vez de funciones*”. Por supuesto, este no es un teorema pues resulta ser que hay muchas cuantizaciones. Pero es una idea que puede ser muy útil, aún más útil que varios teoremas.

6. Regresando al Mundo Cuántico

Comparando la forma canónica de una función simple (variable aleatoria) con la forma canónica de una matriz autoadjunta, podemos dar por analogía una interpretación física en la mecánica cuántica de una

matriz autoadjunta $A = A^*$. Interpretamos cada eigenvalor λ_j de A como un *resultado posible* de un experimento que mide la cantidad física correspondiente a A y no hay otros resultados posibles.

La proyección ortogonal E_j , asociada al eigenvalor λ_j , debe ser interpretada como el evento cuántico correspondiente a la medición de λ_j . Por lo tanto, definimos un *evento (cuántico)* como una proyección ortogonal E en $\text{MAT}(n; \mathbb{C})$, o más bien, un subespacio V de \mathbb{C}^n .

Ejemplo: Cada una de las matrices autoadjuntas $\frac{1}{2}\sigma_1$, $\frac{1}{2}\sigma_2$ y $\frac{1}{2}\sigma_3$ corresponden a una cantidad física. Estas cantidades físicas son las tres componentes (en direcciones ortogonales en el espacio euclideo \mathbb{R}^3) del *spin* de un sistema físico con spin $1/2$. Aquí hablamos solamente de la *matemática* del spin. Para aprender la *física* de spin, hay que leer algún texto sobre física cuántica, como por ejemplo [4] o [19]. Por otro lado, el lector debe darse cuenta que hay muchísimos otros libros excelentes y puede escoger otro según su preferencia.

Ejercicio: Encontrar la resolución espectral para estas tres matrices. ¿Cuáles son los resultados posibles de cada experimento y cuáles sus correspondientes eventos cuánticos?

Si E es un evento y se tiene $\dim_{\mathbb{C}} \text{Ran}(E) = 1$, se dice que es un *estado (puro)* o un *estado cuántico (puro)*. Equivalentemente, un estado es un subespacio $V \subset \mathbb{C}^n$ de dimensión uno. (El porqué de esta definición se encuentra en la definición de evolución temporal de un sistema cuántico, cosa que no explicaremos aquí.)

En física se le llama estado a un vector $z \in \mathbb{C}^n$ con $\|z\| = 1$, pues este vector define al subespacio

$$\mathbb{C}z := \{\lambda z \mid \lambda \in \mathbb{C}\} \subset \mathbb{C}^n,$$

de dimensión uno. Esta *identificación de vectores con espacios* se puede hacer notando que los vectores w y $z \in \mathbb{C}^n$ con $\|w\| = \|z\| = 1$ definen el mismo subespacio $\mathbb{C}w = \mathbb{C}z$ si y sólo si existe $\alpha \in \mathbb{C}$ (con $|\alpha| = 1$) tal que $w = \alpha z$.

Los expertos notarán que el conjunto de los estados cuánticos puros es el llamado *espacio proyectivo complejo* de dimensión $n - 1$ sobre los complejos, denotado por $\mathbb{C}P^{n-1}$. Para los que todavía no son expertos recomendamos nuevamente la *lectura* de estos temas.

Con esto, ya tenemos el lenguaje suficiente para describir la Probabilidad Cuántica.

Un Principio de la Mecánica Cuántica:

Sea A una matriz autoadjunta y supongamos que

$$A = \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_k E_k$$

es su resolución espectral. Si un sistema cuántico empieza en un estado $z \in \mathbb{C}^n$ con $\|z\| = 1$ y se mide la cantidad física asociada a A , entonces el experimento da el valor λ_j con *frecuencia relativa*

$$\langle z, E_j z \rangle = \|E_j z\|^2.$$

(Después de la medición de λ_j , el estado final es $E_j z / \|E_j z\|$, un cambio al que se le llama *el colapso de la función de onda*. No usaremos este hecho.) ■

Se puede demostrar que los números $\langle z, E_j z \rangle$ para $j = 1, \dots, k$ satisfacen las propiedades de frecuencias relativas; es decir,

$$0 \leq \langle z, E_j z \rangle \leq 1 \quad \text{y} \quad \sum_{j=1}^k \langle z, E_j z \rangle = 1.$$

Si se mide λ_j se dice que el evento cuántico E_j ha *sucedido*.

Como cada evento cuántico E se encuentra en la resolución espectral de alguna matriz autoadjunta (por ejemplo de la misma matriz E) debemos definir la probabilidad para cada evento, no solamente para aquellos eventos en la resolución espectral de A .

Por esta razón, definimos la *Probabilidad Cuántica* de que suceda el evento E (arbitrario) dado el estado inicial $z \in \mathbb{C}^n$ (con $\|z\| = 1$) por la expresión

$$\text{Prob}(E; z) := \langle z, E z \rangle = \|E z\|^2.$$

Es importantísimo notar que el mapeo $E \rightarrow \langle z, E z \rangle \in [0, 1]$ (con el estado z fijo) *no* es una función de probabilidad clásica (de Kolmogorov) si $n \geq 2$.

Ahora, supongamos que tenemos una matriz autoadjunta A con resolución espectral dada por

$$A = \sum_{j=1}^k \lambda_j E_j$$

y sea z un estado cuántico, es decir $z \in \mathbb{C}^n$ y $\|z\| = 1$. Entonces, el *valor esperado de A en el estado z* está definido por la suma (sobre j) del valor de A en el evento E_j por la probabilidad del evento E_j en el estado z , es decir,

$$\langle A \rangle_z := \sum_{j=1}^k \lambda_j \text{Prob}(E_j; z) = \sum_{j=1}^k \lambda_j \langle z, E_j z \rangle = \langle z, \sum_{j=1}^k \lambda_j E_j z \rangle = \langle z, A z \rangle.$$

Esta definición nos permite generalizar el concepto de estado. Definimos el mapeo lineal $\rho_z = \rho : \mathcal{A} \rightarrow \mathbb{C}$ donde $\mathcal{A} = \text{MAT}(n; \mathbb{C})$, con un estado z fijo, por la expresión

$$\rho_z(M) = \rho(M) := \langle z, Mz \rangle \quad \text{para } M \in \mathcal{A}.$$

Se tienen las siguientes dos propiedades:

$$\rho \text{ es } \textit{positivo}: \rho(M^*M) \geq 0 \text{ para toda } M \in \mathcal{A}.$$

$$\rho \text{ está } \textit{normalizado}: \rho(I) = 1.$$

Extenderemos el concepto de estado para álgebras-*. Un álgebra-* \mathcal{A} es un álgebra sobre \mathbb{C} , con unidad I , con un mapeo que manda cada $A \in \mathcal{A}$ a un elemento $A^* \in \mathcal{A}$. Para A y $B \in \mathcal{A}$ y para $\lambda \in \mathbb{C}$ este mapeo satisface que

1. $(A + B)^* = A^* + B^*$,
2. $(\lambda A)^* = \lambda^* A^*$,
3. $(AB)^* = B^* A^*$,
4. $A^{**} = A$.

A los elementos $A \in \mathcal{A}$ tales que $A = A^*$ se les llama *Hermitianos* (o *autoadjuntos*). Estos elementos corresponden a las matrices autoadjuntas en la mecánica cuántica y, como veremos más adelante, a las variables aleatorias en la probabilidad clásica. (A veces se dice que cada elemento $A \in \mathcal{A}$ es una variable aleatoria.)

Un *estado* de un álgebra-* \mathcal{A} sobre \mathbb{C} (digamos de dimensión finita) es una función lineal $\rho : \mathcal{A} \rightarrow \mathbb{C}$ que es positiva y está normalizada (como se definió antes).

Un ejemplo de un estado es el mapeo $\rho_T : \text{MAT}(n; \mathbb{C}) \rightarrow \mathbb{C}$ definido como $\rho_T(M) := \text{Tr}(TM)$, donde M y $T \in \text{MAT}(n; \mathbb{C})$, la matriz T es autoadjunta, tiene eigenvalores no negativos y se cumple que $\text{Tr}(T) = 1$. (Aquí Tr denota la traza de una matriz.) Decimos que ρ_T es un *estado mezclado*.

Dejamos como ejercicios el hecho que ρ_T es un estado y que para cada estado $z \in \mathbb{C}^n$ existe una matriz T con las propiedades indicadas arriba tal que $\rho_z = \rho_T$. Además el lector debe verificar que, para $n \geq 2$, existen matrices T (con las propiedades indicadas arriba) tales que para cada estado $z \in \mathbb{C}^n$ se tiene que $\rho_T \neq \rho_z$.

La probabilidad clásica (Ω, P) es un caso particular de los conceptos anteriores: obsérvese que $\mathcal{A} = \{Y \mid Y : \Omega \rightarrow \mathbb{C}\}$ es un álgebra-* sobre

\mathbb{C} y que $\rho(Y) = \sum_{\omega \in \Omega} P(\omega)Y(\omega)$ es un estado. También nótese que los elementos autoadjuntos en \mathcal{A} son las variables aleatorias. Se suele decir que P es el estado, debido que ρ determina a P únicamente.

7. Las Estructuras Básicas

De las secciones anteriores se sigue que la *Probabilidad Cuántica* (en nuestro caso de dimensión finita), que empezó con los trabajos de Murray y von Neumann en la década de los 1930, consta de las siguientes estructuras básicas:

- El espacio vectorial \mathbb{C}^n con su producto interior.
- Los eventos cuánticos.
- Los estados cuánticos.
- Las matrices autoadjuntas (observables).
- Las fórmulas para calcular las probabilidades cuánticas, los valores esperados, etcétera.

8. El Porvenir

La finalidad de este artículo es que al lector le den ganas de seguir estudiando la probabilidad cuántica y otros temas relacionados. Damos ahora unas referencias, pero es una lista pequeña. Son nuestras preferencias. Hay muchas otras buenas referencias.

Antes que nada vale la pena aprender un poco de la física cuántica. Una introducción muy intuitiva se da en [4]. El texto [19] es para *estudiantes* de matemáticas, aunque su subtítulo diga que es para *matemáticos*.

Algunas referencias para estudiar la probabilidad cuántica son los libros [3], [8] y [11] y el artículo [16]. Esta última referencia muestra claramente que hay muchas probabilidades no clásicas.

Después de esto hay que estudiar la teoría espectral en dimensión infinita. Para eso uno puede leer [21] o el volumen I de Reed y Simon [17].

Para estudiar la ecuación de Schrödinger véase [19] si se desea leer algo a nivel introductorio o los volúmenes II, III y IV de Reed y Simon [17] si se desea ver un nivel avanzado. Para estudiar otros modelos

de evolución temporal, incluyendo los procesos cuánticos estocásticos, véase [2].

Para una introducción al spin (y a los bosones y fermiones), véase [19]. Para el cálculo cuántico estocástico, consúltese el texto [13]. Para la probabilidad libre, que es una teoría de probabilidad no clásica muy estudiada, recomendamos [7] y [20].

Por cierto, hay aún más temas interesantes como independencia, álgebras de von Neumann, teoremas de límite central, espacios de Fock, movimiento Browniano, etcétera. Pero el lector debe descubrir el placer de buscar por su propia cuenta en bibliotecas, librerías e internet.

9. Apéndice: Qubits

Un caso interesante de la mecánica cuántica es $n = 2$, donde tenemos que el espacio de estados cuánticos puros es $\mathbb{C}P^1 \cong S^3/S^1 \cong S^2$. (Aquí S^n denota a la esfera de dimensión n de vectores de norma uno en \mathbb{R}^{n+1} .) Un estado tal se llama *qubit* en la *Computación Cuántica* y en la *Teoría de Información Cuántica*. Por lo tanto hay una esfera de qubits.

En el caso $n = 2$ de la probabilidad clásica, tenemos $\Omega = \{\uparrow, \downarrow\}$, digamos. Aquí hay dos estados puros: $\{\uparrow\}$ y $\{\downarrow\}$; cada uno se llama un *bit*. Por lo tanto hay solamente dos bits.

Es por esto que hay una diferencia muy grande entre la *computación cuántica*, basada en qubits, y la *computación clásica*, basada en bits. Por ejemplo, hay *algoritmos cuánticos* muchísimo más rápidos que los correspondientes *algoritmos clásicos* conocidos. Hay mucha investigación en esta área con muchos problemas abiertos. Véase [12], [14], [15] y [18].

10. Agradecimientos

Este artículo empezó como una conferencia que impartí en el evento “Métodos Estocásticos en Sistemas Dinámicos, La Probabilidad y su interacción con otras áreas de la Matemática” que tuvo lugar en el Centro de Investigación en Matemáticas (CIMAT), Guanajuato, Gto., México, 26-30 de enero de 2009. Quiero agradecer a los organizadores (Xavier Gómez Mont, Renato Iturriaga, José Alfredo López Mimbela, Joaquín Ortega y Ekaterina Todorova) por su tan amable invitación para participar en ese evento. Por su ayuda con el uso del español, les doy muchas gracias a Lenin Echavarría Cepeda y un revisor anónimo.

También quiero subrayar que el entusiasmo, interés y apoyo de Luigi Accardi son cosas sin las cuales nunca habría yo empezado a aprender las maravillas de la probabilidad cuántica. Molte grazie, Gigi. Finalmente, les agradezco a los dos árbitros anónimos por sus sugerencias que han mejorado este artículo.

Bibliografía

1. L. Accardi, Quantum probability: an introduction to some basic ideas and trends, en modelos estocásticos, ii, memorias del vi simposio de probabilidad y procesos estocásticos, en *Aportaciones Matemáticas, (Serie Investigación)*, Sociedad Matemática Mexicana, 2001.
2. A. Chebotarev, Lectures on quantum probability, en *Textos Nivel Avanzado 14*, *Aportaciones Matemáticas*, Sociedad Matemática Mexicana, 2000.
3. E. Davies, *Quantum Theory of Open Systems*, Academic Press, 1976.
4. R. Feynman, R. B. Leighton, y M. Sands, *The Feynman Lectures on Physics*, tomo III, Addison Wesley, 1965.
5. J. García-Corte, Una introducción a la probabilidad cuántica, en memorias del xxvi congreso nacional de la smm, en *Aportaciones Matemáticas (Serie Comunicaciones) 14*, Sociedad Matemática Mexicana, 1994.
6. P. Halmos, *Finite-Dimensional Vector Spaces*, Springer, 1974.
7. F. Hiai y D. Petz, The semicircle law, free random variables and entropy, en *Math. Surveys and Monographs, Vol. 77*, American Mathematical Society, 2000.
8. A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, North-Holland, 1982.
9. ———, Statistical structure of quantum theory, en *Lecture Notes in Physics (Monographs)*, Springer, 2001.
10. A. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung, Traducción en inglés: Foundations of the Theory of Probability*, Chelsea, 1956, Springer, 1933.
11. P.-A. Meyer, Quantum probability for probabilists, en *Lecture Notes in Mathematics, vol. 1538*, Springer, 1993.
12. L. Pantaleón-Martínez y R. Quezada, Una introducción a la teoría cuántica de la información, en *Texto para el primer Simposio del Departamento de Matemáticas*, UAM-Iztapalapa, 2008.
13. K. Parthasarathy, *An Introduction to Quantum Stochastic Calculus*, Birkhäuser, 1992.
14. ———, *Lectures on Quantum Computation, Quantum Error Correcting Codes and Quantum Information*, Narosa Publishing House; American Mathematical Society, 2006.

15. D. Petz, *Quantum Information Theory and Quantum Statistics*, Springer, 2007.
16. M. Rédei y S. Summers, Quantum probability theory, preprint, 2006.
17. M. Reed y B. Simon, *Methods of Modern Mathematical Physics, Vols. I–IV*, 19–48, Academic Press, 1972–1979.
18. M. Ruskai, Open problems in quantum information theory, preprint, 2007.
19. A. Sudbery, *Quantum Mechanics and the Particles of Nature: An Outline for Mathematicians*, Cambridge University Press, 1986.
20. D. Voiculescu, K. Dykema, y A. Nica, *Free Random Variables*, American Mathematical Society, 1992.
21. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955.