Introducción a la teoría de probabilidad

pod

Abril - mayo de 2001

Introducción

Este documento presenta brevemente los principios de la *teoría de la probabilidad*. Dicha teoría representa una de las herramientas matemáticas más importantes para la física, en especial para la teoría de la Mecánica Cuántica, así como en los desarrollos de la Física Estadística. La teoría de la probabilidad se presenta en forma de apuntes esquemáticos y sin demostraciones.

3.1. Interpretación de la probabilidad

Probabilidad clásica (a priori): Asigna una probabilidad a un suceso antes de que este ocurra, basándose en el principio de simetría (casos favorables entre casos totales).

Probabilidad frecuencial: La probabilidad de un suceso es la frecuencia con la que se observa.

Probabilidad subjetiva: Se asigna la probabilidad a partir de la información previa.

Probabilidad como lógica: Basada en razonamientos lógicos.

Probabilidad geométrica: Basada en una medida de los sucesos (medida de los sucesos favorables entre medida total).

3.2. Probabilidad axiomática

Definición 3.2.1. (Espacio muestral, E). *Conjunto de resultados posibles, mutuamente excluyentes, de un una variable aleatoria.*

Definición 3.2.2. (Álgebra de sucesos, Ω). *Conjunto de todos los sucesos (subconjuntos) que se pueden formar a partir de E. Si sus elementos son finitos se llama álgebra de sucesos de Boole, si son infinitos pero numerables, se le llama σ-álgebra.*

La definición axiomática de la probabilidad es:

Definición 3.2.3. (Medida de la probabilidad). *A una función*

$$P : \Omega \longrightarrow \mathbb{R} \ .$$

se le llama medida de la probabilidad si cumple las siguientes condiciones:

1. *Si $S \in \Omega$, entonces existe un valor $P(S) \geq 0$, al que llamaremos probabilidad de S.*

2. *La probabilidad del suceso seguro (espacio muestral) es $P(E) = 1$.*

3. *Dada una sucesión numerable de sucesos disjuntos (mutuamente excluyentes dos a dos) $\{S_i\}_{i \in \mathbb{N}}$, entonces:*

$$P\left(\bigcup_i^{\infty} S_i\right) = \sum_i^{\infty} S_i \ .$$

A partir de estos axiomas, se pueden demostrar las siguientes propiedades de la probabilidad.

Teorema 3.2.4. (Probabilidad del suceso imposible). *La probabilidad del suceso imposible (conjunto vacío), es*

$$P(\emptyset) = 0 \ .$$

Teorema 3.2.5. (Suma finita). *Para toda colección finita de sucesos disjuntos $\{S_i\}_{i=1\ldots n}$, se cumple:*

$$P\left(\bigcup_i^{n} S_i\right) = \sum_i^{n} S_i \ .$$

Teorema 3.2.6. (Probabilidad de la unión). *Para todo par de sucesos S_1 y S_2, se cumple:*

$$P(S_1 \cup S_2) = P(S_1) + P(S_2) - P(S_1 \cap S_2) \ .$$

En general, para una colección finita de sucesos $\{S_i\}_{i=1\ldots n}$, se tiene:

$$\begin{array}{rcl} P\left(\bigcup_i^n S_i\right) & = & \sum_i P(S_i) - \sum_{i \neq j} P(S_i \cap S_j) + \sum_{i<j<k} P(S_i \cap S_j \cap S_k) + \\ & & + \cdots + (-1)^{n+1} P\left(\bigcap_i^n S_i\right) \end{array} .$$

Teorema 3.2.7. (Ordenación). *Para todo par de sucesos que cumplen $S \subseteq S'$, entonces, se cumple:*

$$P(S) \leq P(S) \ .$$

Teorema 3.2.8. (Cota). *Para todo suceso $S \in \Omega$, su probabilidad cumple*

$$P(S) \leq 1 \ .$$

3.3. Probabilidad condicionada

La probabilidad de que se verifique un suceso S_1 sabiendo que ha ocurrido un suceso S de llama probabilidad de S_1 condicionada a S, que se define de la siguiente manera.

Definición 3.3.1. (Probabilidad condicionada). *La probabilidad de S_1 condicionado a S, si $P(S) \neq 0$ se define:*

$$P(S_1|S) = \frac{P(S_1 \cap S)}{P(S)} \ .$$

Las principales propiedades de la probabilidad condicionada son:

Teorema 3.3.2. (Probabilidad condicionada). *La probabilidad condicionada, definida de esta manera, cumple los axiomas de probabilidad, y es una medida de la probabilidad del espacio muestral reducido $E_s = E \cap S$.*

Teorema 3.3.3. (Regla de la multiplicación). *Dada una sucesión finita de sucesos $\{S_i\}_{i=1\ldots n}$, se cumple:*

$$P\left(\bigcap_i^n S_i\right) = P(S_1) \cdot P(S_2|S_1) \cdot P(S_3|S_1 \cap S_2) \ldots P(S_n|S_1 \cap \cdots \cap S_{n-1}) \ .$$

Teorema 3.3.4. (Probabilidad total). *Dados un suceso A y una colección finita de sucesos $\{S_i\}_{i=1\ldots n}$ tal que cumplen:*

1. *Mutuamente disjuntos, $S_i \cap S_j = \emptyset \ , \quad \forall i \neq j$*
2. *Recubren el espacio muestral $\bigcup_i S_i = E$*
3. *Tienen partes comunes con A, $A \cap S_i \neq \emptyset \ , \quad \forall i$*

Entonces, se verifica

$$P(A) = \sum_i^n P(A|S_i)P(S_i) \ .$$

El teorema de la probabilidad total proporciona una manera de calcular la contribución de cada una de las causas (S_i) a la probabilidad de la consecuencia (A).

Teorema 3.3.5. (de Bayes o de las hipotesis). *Sea una colección de sucesos $\{S_i\}_{i=1\ldots n}$ que cumplen las condiciones para que el teorema de la probabilidad total se verifique. Entonces,*

$$P(S_i|A) = \frac{P(A|S_i)P(S_i)}{P(A)} = \frac{P(A|S_i)P(S_i)}{\sum P(A|S_i)P(S_i)} \ .$$

Donde $P(S_i|A)$ es la probabilidad a posteriori o hipótesis; y $P(A|S_i)$ es la verosimilitud.

3.4. Independencia de sucesos

Definición 3.4.1. (Dos sucesos independientes). *Dos sucesos S y S_i son independientes si y solo si se cumple*

$$P(S_i \cap S) = P(S_i)P(S) \ .$$

Esta definición no es suficiente si tenemos un mayor número de sucesos.

Definición 3.4.2. (Sucesos mutuamente independientes). *Los sucesos de una colección finita $\{S_i\}_{i=1\ldots n}$ son mutuamente independientes si cumplen*

$$\begin{array}{lcll} P(S_i \cap S_j) & = & P(S_i)P(S_j) & si \quad i \neq j \\ P(S_i \cap S_j \cap S_k) & = & P(S_i)P(S_j)P(S_k) & si \quad i \neq j \neq k \\ & \vdots & & \\ P\left(\bigcap_i^n S_i\right) & = & \prod_i^n P(S_i) & \end{array} \ .$$

3.5. Variable aleatoria o estocástica discreta

Estudiaremos una variable aleatoria X que puede tomar un conjunto de valores $\{x_i\}$ numerable (finito o infinito).

Definición 3.5.1. (Distribución de probabilidad discreta). *La función de distribución de probabilidad de una variable discreta X, asigna a cada valor x_i de la variable la probabilidad del suceso que consiste que la variable tome dicho valor $\{X = x_i\}$:*

$$P(\{X = x_i\}) = p_i \ .$$

La función de distribución de probabilidad discreta debe verificar:

1. *Cota:* $0 \geq p_i \geq 1$
2. *Normalización* $\sum p_i = 1$

Algunas definiciones de utilidad:

Definición 3.5.2. (Valor esperado o media). *El valor esperado, o media, de una distribución se define mediante la expresión:*

$$E(X) = \langle X \rangle = \bar{X} = \sum_i x_i p_i \ .$$

Definición 3.5.3. (Momentos de la distribución). *El momento de orden n de una distribución se define como el valor esperado de X^n:*

$$\langle X^n \rangle = \sum_i x_i^n p_i \ .$$

Definición 3.5.4. (Varianza y desviación típica). *La varianza de una distribución se define:*

$$\mathrm{Var}(X) = \langle (X - E(X))^2 \rangle = \langle X^2 \rangle - \langle X \rangle^2 \ .$$

La desviación típica se define como la raíz cuadrada de la varianza:

$$\sigma_X = \sqrt{\mathrm{Var}(X)} = \sqrt{\langle X^2 \rangle - \langle X \rangle^2} \ .$$

Algunas propiedades del operador valor esperado vienen dados por el siguiente

Teorema 3.5.5. (Propiedades de $E(X)$). *Para una distribución bien definida, el operador de valor esperado cumple:*

1. *Escala:* $E(cX) = cE(X) \ , \quad \forall c \in \mathbb{R}$
2. *Adición:* $E(X \pm Y) = E(X) \pm E(Y)$
3. *Independencia:* $E(XY) = E(X)E(Y)$ *si* X *e* Y *son independientes.*
4. *Composición:* $E[g(x)] = \sum g(x_i)p_i$
5. *No desviado:* $E(X - E(X)) = 0$

Por último, veamos algunas propiedades de la varianza:

Teorema 3.5.6. (Propiedades de la varianza). *Para una distribución bien definida, la varianza cumple:*

1. *Origen:* $\mathrm{Var}(X + c) = \mathrm{Var}(X) \ , \quad \forall c \in \mathbb{R}$
2. *Adición:* $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$ *si* X *e* Y *son independientes.*
3. *Escala:* $\mathrm{Var}(cX) = c^2 \mathrm{Var}(X)$

3.6. Variables aleatorias continuas

Para una variable aleatoria que puede tomar cualquier valor dentro de un rango (que puede ser infinito), se definen las funciones de densidad de probabilidad y ´de distribución acumulada.

Definición 3.6.1. (Función de densidad de probabilidad). *Para una variable aleatoria continua X, el valor $f(x)\mathrm{d}x$ identifica la probabilidad del suceso que se verifica cuando el valor de X está en el intervalo $(x, x + \mathrm{d}x)$, es decir:*

$$f(x)\mathrm{d}x = P(\{x \leq X \leq x + \mathrm{d}x\}) \ .$$

La función de densidad de probabilidad $f(x)$ debe cumplir:

1. *No negatividad:* $f(x) \geq 0 \ , \quad \forall x$

2. *Normalización:*
$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1 \ .$$

3. *Probabilidad:*
$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)\mathrm{d}x \ .$$

Definición 3.6.2. (Función de distribución acumulada). *La función de distribución acumulada $F(x)$ se define como la probabilidad de que la variable X tenga un valor inferior o igual a x. En concreto:*

$$F(x) = P(\{X \leq x\}) = \int_{-\infty}^{x} f(x)\mathrm{d}x \ .$$

La función de distribución acumulada debe cumplir:

1. *Límite inferior:* $F(x \to -\infty) = 0$

2. *Límite superior:* $F(x \to \infty) = 1$

3. *Monotonía:* $F(x)$ *es creciente.*

4. *Probabilidad de un rango:* $P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1)$

Igual que para las distribuciones discretas, podemos definir los momentos de una distribución:

Definición 3.6.3. (Momentos de una distribución continua). *El momento de orden n de una distribución continua se define:*

$$\langle X^n \rangle = \int_{-\infty}^{\infty} x^n f(x)\mathrm{d}x \ .$$

El valor esperado se define como el momento de orden $n = 1$.

La varianza y la desviación tienen la misma definición que en el caso discreto.

Otra definición de interés resulta la función característica de la distribución:

Definición 3.6.4. (Función característica). *La función característica de una distribución se define mediante:*

$$\phi_X(k) = E(\mathrm{e}^{ikX}) = \int_{-\infty}^{\infty} \mathrm{e}^{ikx} f(x)\mathrm{d}x \ .$$

El desarrollo de Taylor de la exponencial muestra que los momentos de cada orden son los coeficientes de la expansión:

$$\phi_X = \sum \frac{(ik)^n}{n!}\langle x^n\rangle \ ; \qquad \langle x^n\rangle = \left.\frac{\mathrm{d}^n\phi_X}{i^n \mathrm{d}k^n}\right)_{k=0} \ .$$

Por la definición de valor esperado, podemos ver que la función característica es la transformada de Fourier de la función de densidad de probabilidad, por tanto, esta última se puede calcular de la primera mediante:

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-ikx}\phi_X(k)\mathrm{d}k \ .$$

3.7. Ejemplos de distribuciones

3.7.1. Distribuciones discretas

3.7.1.1. Pruebas de Bernoulli

Experimento con dos resultados posibles (1 y 0, con probabilidades respectivas p y $q = 1 - p$). La probabilidad de obtener n veces el valor 1 en un orden concreto de N tiradas viene dado por:

$$P = p^n(1-p)^{N-n} \ .$$

3.7.1.2. Distribución binomial

Una variable discreta X sigue la distribución binomial si representa el número de éxitos en una N pruebas de Bernoulli independientes. Enonces, tenemos:

$$P(n) = \frac{N!}{(N-n)!n!}p^n(1-p)^n \ .$$

La distribución normal tiene por valor esperado $E(X) = Np$, su momento de orden 2 es $\langle X^2\rangle = Np(Np+q)$ y su varianza es $\mathrm{Var}(X) = Npq$

3.7.1.3. Distribución geométrica

La distribución geométrica representa la probabilidad de que, en una serie de pruebas de Bernoulli, el primer éxito (1) se obtenga en la n-ésima tirada. La función de distribución es:

$$P(n) = (1-p)^{n-1}p \ .$$

Su valor esperado es $\langle n\rangle = p^{-1}$ y su desviación típica $\sigma = \sqrt{\mathrm{Var}(X)} = p^{-2}(1-p)$

3.7.1.4. Distribución de Poisson

Una variable aleatoria discreta X sigue la distribución de Poisson si la función de distribución es:

$$P(X = n) = \frac{\lambda^n \mathrm{e}^{-\lambda}}{n!} \ .$$

donde λ es el valor esperado y la varianza de la distribución.

3.7.2. Distribuciones continuas. Distribución Gaussiana y normal

La versión discreta de la distribución normal es:

$$P_N(n) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left\{ -\frac{(n-\mu)^2}{2\sigma_N^2} \right\} \ .$$

La variante continua de la distribución Gaussiana, con parámetros μ y σ (valor esperado y desviación típica), viene definida por la función de densidad Gaussiana:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left\{ -\frac{(n-\mu)^2}{2\sigma^2} \right\} \ .$$

Por cambios de variables, se puede reducir a la distribución normal, un caso especial donde $\mu = 0$ y $\sigma = 1$. Los momentos de la distribución normal vienen dados por:

$$\langle X^n \rangle = \begin{cases} 0 & \text{si} \quad n \text{ par} \\ \frac{2^{n/2}}{\sqrt{2}} \Gamma(\frac{n+1}{2}) & \text{si} \quad n \text{ impar} \end{cases} \ .$$