



Clase 4 - Mecánica Estadística

Duración: 1:10 hrs.

Publicada el 13 de septiembre de 2016

Prof. Álvaro Núñez

1. Conceptos Básicos

- Probabilidad de una distribución de probabilidad.
- Entropía: degeneración de un macroestado.
- Ensembles: todas las posibles realizaciones de un evento.

2. Dados cargados

Consideraremos un dado (cubo con los números $\{1, \dots, 6\}$).

- En un dado justo, cada número sale con probabilidad $1/6$.
- El promedio es trivial... $(1 + 2 + 3 + 4 + 5 + 6)/6 = 3,5$.
- Problema inverso... dado el promedio (digamos $\bar{p}$) ¿cuál es la distribución de probabilidad?.
- Hipercubo de dimensión 6, con dos *planos*

$$\sum_i p_i = 1 \quad (1)$$

$$\sum_i i p_i = \bar{p} \quad (2)$$

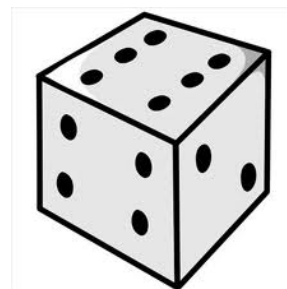


Figura 1: Dado

... no alcanzan los datos!. Hay muchas distribuciones de probabilidad que tienen el promedio requerido. ¿que hacemos?

¡A veces es necesario tomar decisiones sin toda la información necesaria! Si tuvieramos que *apostar* por la distribución de probabilidades más probable... ¿que contestaríamos?.

3. Ensembles y entropía

- Nuestra solución parte considerando el conjunto de todos los posibles resultados (microestados) de tirar un dado N veces. El conjunto tiene 6^N elementos.
- Digamos que de los N tiros, N_1 resulto en 1, ..., y N_6 resulto en 6.

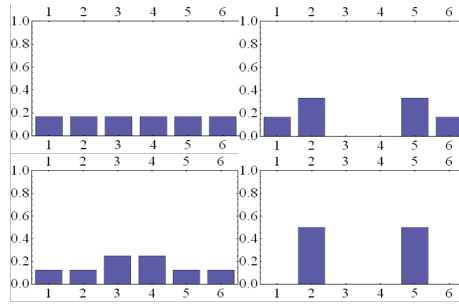


Figura 2: Ejemplos de distribuciones de probabilidad que dan promedio 3.5.

- $(N_1, \dots, N_6) \sim \text{macroestado}$.
- Si N es muy grande esperamos $N_i \sim N p_i$.
- De los 6^N posibles resultados hay,

$$\Omega = \frac{N!}{N_1! N_2! N_3! N_4! N_5! N_6!} \quad (3)$$

microestados que son consistentes con el macroestado $(N_1, \dots, N_6)$.

- La distribución más probable, es aquella que es consistente con el mayor número de microestados.

4. Solución: la distribución de probabilidad más probable

- Problema matemático: maximizar la función Ω , sujeto a la restricción $\sum_i p_i = 1$. Es más conveniente maximizar $S = \log \Omega$.
- La función S se conoce como la entropía de un macroestado. Esencialmente mide su degeneración: el número de microestados consistentes con él.
- Las matemáticas son fáciles (multiplicadores de Lagrange):

$$s(\{p_i\}) = \frac{S}{N} = -(p_1 \log p_1 + \dots + p_6 \log p_6) \quad (4)$$

- Obtenemos que la distribución más probable es la distribución uniforme: $p_i = 1/6$.
- Si ahora fijamos el promedio el cálculo se hace algo más interesante. Debemos maximizar la función:

$$f(\{p_i\}, \lambda, \beta) = s(\{p_i\}) - \lambda(\sum_i p_i - 1) - \beta(\sum_i i p_i - \bar{p}). \quad (5)$$

- Optimizando respecto a $\{p_i\}$, obtenemos:

$$p_i = e^{1+\lambda} e^{i\beta} \quad (6)$$

- La normalización de las probabilidades nos permite obtener λ :

$$e^{1+\lambda} = \frac{1}{Z} = e^{-\beta} \frac{e^{\beta} - 1}{e^{6\beta} - 1} \quad (7)$$

- Del mismo modo obtenemos:

$$\frac{1}{Z} \frac{\partial Z}{\partial \beta} = \frac{\partial}{\partial \beta} \log Z = \bar{p} \quad (8)$$

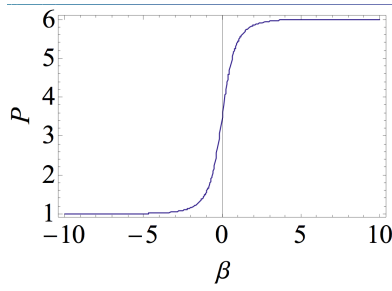


Figura 3: Promedio en función de β . Para obtener un promedio dado debemos invertir esta relación.

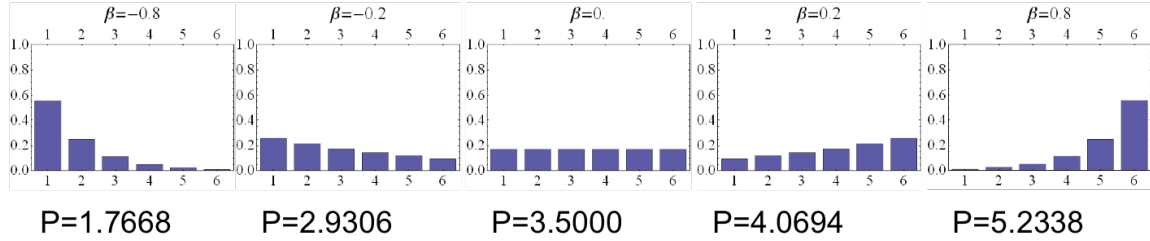


Figura 4: Distribuciones más probables para distintos promedios (distintos β).

5. Entropía de una distribución de probabilidad

$$S = \log \Omega = - \sum_i p_i \log p_i$$

- La definimos así, inspirados en el problema del dado cargado.
- Representa el logaritmo del número de estados microscópicos distintos consistentes con la distribución.
- El logaritmo del número de secuencias de resultados que entregan las probabilidades a posteriori igual a la distribución.

¡A veces es necesario tomar decisiones sin toda la información necesaria! Si tuviéramos que *apostar* por la distribución de probabilidades más probable, lo lógico es maximizar su entropía.

6. Ejemplo: entropía de un evento binario.

- Consideremos un problema con un sistema binario, con estados 0 y 1.
- Los estados se dan con probabilidad p_0 y p_1 .
- La entropía es por definición:

$$S = -(p_0 \log p_0 + p_1 \log p_1) \quad (9)$$

- Naturalmente las probabilidades deben sumar 1, $p_0 + p_1 = 1$.
- La función, depende entonces de una sola variable:

$$S = -(p_0 \log p_0 + (1 - p_0) \log(1 - p_0)) \quad (10)$$

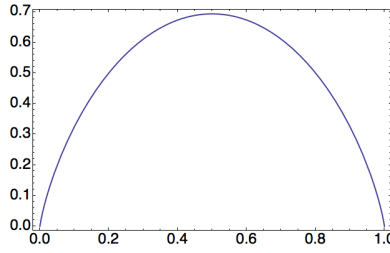


Figura 5: Entropía de un sistema con dos estados en función de la probabilidad de un estado.

7. Algunas propiedades matemáticas de la entropía

Ahora consideraremos a la entropía como una función de s variables donde s es el número de estados.

$$\mathbf{S} = \mathbf{S}(p_1, \dots, p_s) \quad (11)$$

- Es una función simétrica. $\mathbf{S}(p_1, p_2, \dots, p_s) = \mathbf{S}(p_2, p_1, \dots, p_s)$, etc.
- Es máxima para el caso de equiprobabilidad.

$$\mathbf{S}(p_1, \dots, p_s) \leq \mathbf{S}(1/s, \dots, 1/s) \quad (12)$$

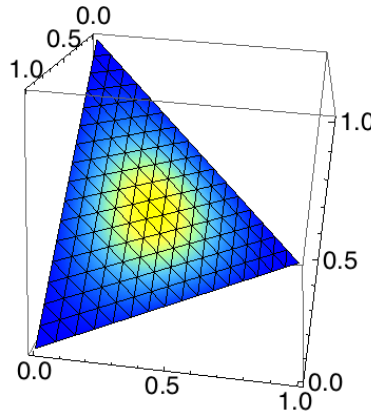


Figura 6: Entropía de un sistema con tres estados en función de la probabilidad de cada estado.

- La entropía de eventos independientes es aditiva. Supongamos que tenemos dos dados, cada uno cargado de manera distinta, con probabilidades para cada número i dadas por $p_i^{(1)}$ y $p_i^{(2)}$ respectivamente. En cada tiro obtenemos dos números (i, j) asociados al primer y segundo dado respectivamente. Tenemos 36 posibilidades dadas por todos los pares (i, j) . La entropía de los dos dados es, por definición:

$$\mathbf{S}^{\text{dos dados}} = - \sum_{(i,j)} p_{(i,j)} \log p_{(i,j)} \quad (13)$$

Como los resultados de los dados son independientes tenemos: $p_{(i,j)} = p_i^{(1)} \times p_j^{(2)}$ de donde obtenemos:

$$\mathbf{S}^{\text{dos dados}} = \mathbf{S}^{(1)} + \mathbf{S}^{(2)} \quad (14)$$

8. Ejemplo: el alfabeto

La cantidad de información asociada a un mensaje es un concepto difícil de precisar pues tiene muchas interpretaciones. En ciertos contextos restringidos resulta conveniente definir información de la siguiente manera. Consideremos un mensaje de N caracteres tomados de algún alfabeto con m letras. Por ejemplo, usando el alfabeto $\mathcal{A} = \{ @, \sqrt{}, \triangle, \square \}$, podemos escribir (y enviar a alguien) el siguiente mensaje:

$$\mathcal{M}_1 = @@ \square \sqrt{} \sqrt{} @ @ @ \square \triangle @ @ @ \quad (15)$$

$\mathcal{M}_1$ tiene 15 caracteres. Existen en realidad $4^{15} = 1,073,741,824$ (en general m^N) posibles mensajes de dicha longitud usando el alfabeto $\mathcal{A}$. Los siguientes son otros posibles mensajes:

$$\mathcal{M}_2 = @ @ @ @ @ @ @ @ @ @ @ @ @ @ @ @ @ \quad (16)$$

$$\mathcal{M}_3 = \square @ \square @ \triangle \square @ \square @ \triangle \square @ \square @ \triangle \quad (17)$$

todos ellos igualmente interesantes. Si nos llega un mensaje lo único que objetivamente recibimos es un número entre 1 y m^N . Sin embargo las letras no se repiten todas uniformemente. No todos los números de 1 a m^N son mensajes permitidos, hay diversas de restricciones gramaticales. En logaritmo del número de mensajes por caracter consistente con la distribución de probabilidad de caracteres es la entropía del lenguaje.

9. Problemas

- Considere el problema de tirar una moneda al aire. Suponga que el resultado es cara con probabilidad p y sello con probabilidad $1 - p$. Encuentre la entropía y demuestre que es máxima cuando $p = 1/2$.
- La traducción de los mensajes del material genético en proteínas se inicia mediante “start codons” en la secuencia de ARN (ácido ribonucleico). Usualmente, en organismos simples, la palabra clave usada como start codon es la secuencia “AUG”. Supongamos que el ARN es una cadena simple y consiste en una secuencia de bases, cada una descrito por una sola letra. El alfabeto de letras de ARN consiste en “A”, “C”, “U”, “G”. Considerando, por simplicidad, que las letras son no correlacionadas y equiprobables: ¿Cuál es la probabilidad de tener una secuencia “AUG” dentro de una palabra de seis letras? ¿Cuál es la entropía del código genético en esta aproximación?
- Considere una idealización de un crystal que tiene N sitios y el mismo número de sitios intersticiales (lugares intermedios entre sitios del crystal que pueden albergar átomos). Sea ε es la energía necesaria para mover un átomo desde el crystal a un sitio intersticial y n el número de átomos ocupando sitios intersticiales. Determine la entropía en función de la energía total $E = n\varepsilon$. Asuma $1 \ll n \ll N$