Repaso de Probabilidad Discreta

Alan Reyes-Figueroa

Criptografía y Cifrado de Información (Aula 03) 20.julio.2021

La criptografía moderna se desarrolló como una ciencia rigurosa donde las construcciones siempre van acompañadas de una prueba de seguridad. El lenguaje utilizado para describir la seguridad se basa en los fundamentos de probabilidad discreta.

Definición

Sea Ω un conjunto finito. Una distribución de probabilidad p sobre Ω, es una función, $p : \Omega \to [0, 1]$ tal que

$$\sum_{\mathbf{x} \in \Omega} p(\mathbf{x}) = 1.$$

Al valor $p(\mathbf{x})$ se les llama el **peso** o **probabilidad** de $\mathbf{x}$.

Ejemplo: Letras en el español. En este caso $\Omega = \{\mathbf{a}, \mathbf{b}, \mathbf{c}, , \mathbf{e}, \ldots, \mathbf{x}, \mathbf{y}, \mathbf{z}\}$. Un ejemplo de distribución de probabilidad sobre Ω es la distribución de ocurrencia de las letras en textos escritos. Aquí

$$p(\mathbf{e}) = 0.1218, \; p(\mathbf{a}) = 0.1152, \; p(\mathbf{o}) = 0.0868, \; p(\mathbf{s}) = 0.0798, \ldots$$

es una distribución de probabilidad.

Ejemplo: Cadenas de n-bits. Consideramos el conjunto $\Omega = \{0,1\}^n = \{\mathbf{w} : \mathbf{w} \text{ es una cadena de } n \text{ bits}\}$.
Por ejemplo, si $\Omega = \{0,1\}^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$

$\mathbf{x}$	000	001	010	011	100	101	110	111
$p(\mathbf{x})$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{16}$	$\frac{1}{8}$	0	$\frac{1}{16}$	0

y p es una distribución de probabilidad.

Existen muchas distribuciones discretas de probabilidad. Entre las más comunes tenemos:

Ejemplos:

- Distribución uniforme: Si $|\Omega| = n$, para todo $\mathbf{x} \in \Omega$, $p(\mathbf{x}) = \frac{1}{|\Omega|} = \frac{1}{n}$.
- Distribución puntual en $\mathbf{x}_0$: En este caso, $p(\mathbf{x}_0) = 1$, mientras que $p(\mathbf{x}) = 0$, para todo $\mathbf{x} \neq \mathbf{x}_0$.

Algo que será útil es escribir las probabilidades como un vector. Si $\Omega = \{\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n\}$, podemos representar su distribución de probabilidad por el vector $(p_1, p_2, \ldots, p_n) \in \mathbb{R}^n$. Aquí

$$p_i = p(\mathbf{x}_i), \qquad \text{para } i = 1, 2, \ldots, n.$$

Ej: $\big(p(000), p(001), p(010), p(011), p(100), p(101), p(110), p(111)\big)$.

Definición

*Un **evento** en el espacio de probabilidad Ω es cualquier subconjunto $A \subseteq \Omega$. Definimos la probabilidad de un evento como*

$$p(A) = \sum_{\mathbf{x} \in A} p(\mathbf{x}).$$

Las distribuciones de probabilidad cumplen varias propiedades:

- $p(\varnothing) = 0$, $p(\Omega) = 1$,
- $p(A^c) = 1 - p(A)$,
- $p(A \cup B) = p(A) + p(B)$, si A y B son **eventos disjuntos** ($A \cap B = \varnothing$),
- En general, $p(A \cup B) = p(A) + p(B) - p(A \cap B)$,
- Cuando la distribución es la uniforme, $p(A) = \frac{|A|}{|\Omega|}$.

Ejemplo: $\Omega = \{0,1\}^3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$ y

$\mathbf{x}$	000	001	010	011	100	101	110	111
$p(\mathbf{x})$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{16}$	$\frac{1}{8}$	0	$\frac{1}{16}$	0

Hallar la probabilidad de los siguientes eventos:

- $A = \{$cadenas que terminan en 10$\}$: $\quad p(A) = \frac{9}{16}$.
- $B = \{$cadenas cuyo segundo bit es 0$\}$: $\quad p(B) = \frac{3}{8}$.

Ejemplo: $\Omega = \{0,1\}^8$, con la distribución uniforme ($p(\mathbf{x}) = \frac{1}{2^8} = \frac{1}{256}, \forall \mathbf{x}$). ¿Cuál es la probabilidad de los siguientes?

- $A = \{$ la cadenas que terminan en 11$\}$: $\quad p(A) = \frac{1}{4}$.
- $B = \{$ la cadenas que terminan en 100$\}$: $\quad p(B) = \frac{1}{8}$.

Probabilidad Discreta

Definición

Para una cadena $\mathbf{x} \in \{0,1\}^n$*, definimos*

- $\textbf{\textit{lsb}}_k(\mathbf{x})$ *= sus k bits menos significativos.*
- $\textbf{\textit{msb}}_k(\mathbf{x})$ *= sus k bits más significativos.*

Ejemplo: $\mathbf{lsb}_4(011100110101) = 0101, \quad \mathbf{msb}_5(011100110101) = 01110.$

Si $\Omega = \{0,1\}^n$, con la distribución uniforme, se tiene

- $\mathbb{P}(\{\mathbf{lsb}_k(\mathbf{x}) = x_k \cdots x_1\}) = \frac{1}{2^k}$.
- $\mathbb{P}(\{\mathbf{lsb}_k(\mathbf{x}) = x_n \cdots x_{n-k+1}\}) = \frac{1}{2^k}$.
- Cualquier evento que coniste en fijar k bits tiene probabilidad $\mathbb{P} = \frac{1}{2^k}$.

¿Cómo calcular o acotar la probabilidad de la unión?

Propiedad (Cota de la Unión)

Para cualesquiera dos eventos $A, B \subseteq \Omega$,

- $p(A \cup B) \leq p(A) + p(B)$,
- $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.

Propiedad

Para cualesquiera tres eventos $A, B, C \subseteq \Omega$,

- $p(A \cup B \cup C) \leq p(A) + p(B) + p(C)$,
- $p(A \cup B \cup C) = p(A) + p(B) + p(C) - p(A \cap B) - p(A \cap C) - p(B \cap C) + p(A \cap B \cap C)$.
  (principio de inclusión-exclusión)

En general, $p\left(\bigcup_{i=1}^{k} A_i\right) \leq \sum_{i=1}^{k} p(A_i)$.

Ejemplo: $\Omega = \{0,1\}^{16}$. Dar una cota para la probabilidad del evento

$$A = \{\mathbf{x} \in \Omega : \ \mathbf{lsb}_3(\mathbf{x}) = 111 \text{ ó } \mathbf{msb}_3(\mathbf{x}) = 111\}.$$

Podemos escribir A como unión de dos eventos, $A = A_1 \cup A_2$, donde $A_1 = \{\mathbf{x} \in \Omega : \ \mathbf{lsb}_3(\mathbf{x}) = 111\}$ y $A_2 = \{\mathbf{x} \in \Omega : \ \mathbf{msb}_3(\mathbf{x}) = 111\}$.

Sabemos $p(A_1) = \frac{1}{8}$, $p(A_2) = \frac{1}{8}$. Entonces

$$p(A) = p(A_1 \cup A_2) \leq p(A_1) + p(A_2) = \tfrac{1}{8} + \tfrac{1}{8} = \tfrac{1}{4}.$$

Si queremos el valor exacto

$$p(A) = p(A_1 \cup A_2) = p(A_1) + p(A_2) - p(A_1 \cap A_2) = \tfrac{1}{8} + \tfrac{1}{8} - \tfrac{1}{64} = \tfrac{15}{64}.$$

Definición

Sea (Ω, p) un espacio de probabilidad discreto. Una ***variable aleatoria (discreta)*** *es una función $X : \Omega \to V$, a un conjunto de valores V (e.g. $V = \mathbb{R}$, $V = \mathbb{R}^n$, $V = \mathbb{N}$, ó $V = \{0, 1\}$).*

El objetivo de las variables aleatorias es traspasar la distribución p en el conjunto Ω, a un ambiente donde sea más fácil medir cantidades: $\mathbb{R}$.

Ejemplo: $\Omega = \{0, 1\}^n$, p la distribución uniforme. Considere

$$X : \Omega \to \mathbb{R}, \quad X(\mathbf{x}) = \mathbf{lsb}_1(\mathbf{x}).$$

X codifica el valor del último bit y toma valores 0 ó 1. En particular

$$\mathbb{P}(X = 0) = \mathbb{P}(\{\mathbf{x} : \mathbf{lsb}_1(\mathbf{x}) = 0\}) = \tfrac{1}{2}, \qquad \mathbb{P}(X = 1) = \mathbb{P}(\{\mathbf{x} : \mathbf{lsb}_1(\mathbf{x}) = 1\}) = \tfrac{1}{2}.$$

Ejemplo: Suma de los bits en una cadena.

$\Omega = \{0, 1\}^2$, p la distribución uniforme. $X_1 = \textbf{msb}_1(\mathbf{x})$, $X_2 = \textbf{lsb}_1(\mathbf{x})$. Tenemos que

$$\mathbb{P}(X_1 = 0) = \tfrac{1}{2}, \quad \mathbb{P}(X_1 = 1) = \tfrac{1}{2}, \quad \mathbb{P}(X_2 = 0) = \tfrac{1}{2}, \quad \mathbb{P}(X_2 = 1) = \tfrac{1}{2}.$$

Definimos la variable aleatoria $Z = X_1 + X_2$ (suma de bits).

$X_1 \setminus X_2$	0	1
0	0	1
1	1	2

La distribución de Z es: $\mathbb{P}(Z = 0) = \tfrac{1}{4}, \quad \mathbb{P}(Z = 1) = \tfrac{1}{2}, \quad \mathbb{P}(Z = 2) = \tfrac{1}{4}.$

Variables Aleatorias

Ejercicio: ¿Cuál es la distribución de la suma de bits para cadenas binarias de longitud n?

- Hacer el cálculo para $n = 3$, $n = 4$, $n = 5$ y tratar de hallar un patrón general.
- ¿Les recuerda a algo las probabilidades que se obtienen?