CI2612: Algoritmos y Estructuras de Datos II

Blai Bonet

Universidad Simón Bolívar, Caracas, Venezuela

Análisis probabilístico

Objetivos

- Estudiar los elementos básicos de la teoría de probabilidad para realizar los análisis necesarios de algoritmos

- Ejemplos de análisis probabilístico

Espacio de probabilidad

Las probabilidades se definen sobre un **espacio de probabilidad**

El espacio de probabilidad lo conforman:

– el **espacio muestral** Ω que contiene todos los **posibles resultados** del experimento o proceso probabilístico

– el conjunto $\mathcal{F}$ de **eventos a considerar** (en nuestro caso todos los subconjuntos de Ω)

– la **función o medida de probabilidad** $\mathbb{P}(\cdot)$ definida sobre los eventos

Ejemplo: Lanzar una moneda

Considere el experimento de **lanzar una moneda**

Existen dos posibles resultados del experimento: $\Omega = \{H, T\}$

El conjunto de eventos es $\mathcal{F} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$:

– El evento $\emptyset$ denota que el experimento no dió ninguń resultado

– El evento $\{H\}$ denota que la moneda salió cara

– El evento $\{T\}$ denota que la moneda salió cruz

– El evento $\{H, T\}$ denota que la moneda salió cara o cruz

Si la moneda es insesgada (resultados equiprobables), definimos:
$\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\{H\}) = \mathbb{P}(\{T\}) = \frac{1}{2}$, y $\mathbb{P}(\{H, T\}) = 1$

Eventos

Para cada $\omega \in \Omega$, el evento $\{\omega\}$ se llama **evento atómico** y representa un único resultado del experimento

Los otros eventos denotan conjuntos de **posibles resultados** los cuales son útiles para el análisis del experimento

Por ejemplo, si el experimento es lanzar un dado, podemos hablar del evento que el dado cae en un número par

Después de realizar el experimento podemos decir cuando un evento es cierto o falso. Si el resultado del experimento es ω, decimos que el evento E es cierto si $\omega \in E$. Si $\omega \notin E$, decimos que E es falso

Función (medida) de probabilidad

La función de probabilidad $\mathbb{P}$ es una función $\mathbb{P} : \mathcal{F} \to [0, 1]$ que asigna probabilidades a los eventos y debe satisfacer:

P1. $\mathbb{P}(\Omega) = 1$

P2. para toda colección $\{E_i\}_i$ de eventos que sean **disjuntos dos a dos** (i.e. $E_i \cap E_j = \emptyset$ para todo $1 \leq i < j \leq n$):

$$\mathbb{P}(\cup_{i=1}^n E_i) \;=\; \mathbb{P}(E_1) + \mathbb{P}(E_2) + \cdots + \mathbb{P}(E_n)$$

Por lo tanto, $1 = \mathbb{P}(\Omega) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset) \implies \mathbb{P}(\emptyset) = 0$

Probabilidad del evento complemento

Considere un evento E que denota un posible resultado del experimento

$\mathbb{P}(E)$ es la probabilidad de que dicho resultado suceda mientras que $\mathbb{P}(E^c)$, donde $E^c = \Omega \setminus E$, es la probabilidad de que no suceda

Por la aditividad de la medida y la propiedad P1:

$$\mathbb{P}(E^c) \;=\; \mathbb{P}(\Omega) - \mathbb{P}(E) \;=\; 1 - \mathbb{P}(E)$$

Ejemplo: Lanzar dos monedas

Ahora considere dos monedas **“insesgadas e independientes”**

Definimos:

- $\Omega = \{HH, HT, TH, TT\}$
- $\mathcal{F} = \{E \subseteq \Omega\}$
- $\mathbb{P}(E) = \frac{1}{4}$ para todo evento atómico E (la probabilidad de los demás eventos queda determinada por la propiedad P2)

Probabilidad **primera moneda** salga cara: $\mathbb{P}(\{HH, HT\}) = \frac{1}{2}$

Probabilidad **segunda moneda** salga cara: $\mathbb{P}(\{HH, TH\}) = \frac{1}{2}$

Independencia de dos eventos

Considere un espacio $(\Omega, \mathcal{F}, \mathbb{P})$ y dos eventos $E_1, E_2 \in \mathcal{F}$

E_1 es **independiente** de E_2 si conocer alguna información sobre la ocurrencia de E_1 no altera la probabilidad de la ocurrencia de E_2 (y vice versa)

Formalmente, E_1 es independiente de E_2 ssi $\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_1) \times \mathbb{P}(E_2)$

En el ejemplo, considere $E_1 = \{HH, HT\}$ y $E_2 = \{HH, TH\}$ que denotan los eventos en donde la primera y segunda moneda salen cara respectivamente:

$$\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(\{HH\}) = \tfrac{1}{4} = \tfrac{1}{2} \times \tfrac{1}{2} = \mathbb{P}(E_1) \times \mathbb{P}(E_2)$$

Independencia de más de dos eventos

Dos formas de definirla

Considere la colección $\mathcal{E} = \{E_1, E_2, \ldots, E_n\}$ de eventos:

- $\mathcal{E}$ es **independiente dos a dos** ssi E_i es independiente de E_j para cualquier (i, j) con $i \neq j$; i.e. $\mathbb{P}(E_i \cap E_j) = \mathbb{P}(E_i) \times \mathbb{P}(E_j)$ para todo $1 \leq i < j \leq n$
- $\mathcal{E}$ es **mutuamente independiente** ssi $\mathbb{P}(\cap_{E \in \mathcal{A}} E) = \prod_{E \in \mathcal{A}} \mathbb{P}(E)$ para todo subconjunto $\mathcal{A} \subseteq \mathcal{E}$ de eventos

Las definiciones no son equivalentes, pero la segunda implica la primera

Probabilidad condicional

Considere el experimento de lanzar un dado insesgado y los eventos $A =$ “el dado cae en múltiplo de 3 ó 5” y $B =$ “el dado cae par”

Claramente: $\mathbb{P}(A) = \frac{3}{6} = \frac{1}{2}$ y $\mathbb{P}(B) = \frac{3}{6} = \frac{1}{2}$ pero $\mathbb{P}(A \cap B) = \frac{1}{6}$

¿Cuál es la probabilidad del evento A si sabemos que B es **cierto**?

De las tres posibilidades para el evento B (i.e. el dado sale 2, 4, o 6), solo una hace cierta al evento A. Por lo tanto, la respuesta es $\frac{1}{3}$

Definimos la **probabilidad condicional** $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$

[por lo general denotamos $\mathbb{P}(A \cap B)$ con $\mathbb{P}(A, B)$]

Independencia y probabilidad condicional

Considere dos eventos independientes A y B:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A,B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A) \times \mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A)$$

Ahora considere dos eventos A y B tales que $\mathbb{P}(A|B) = \mathbb{P}(A)$

$$\mathbb{P}(A,B) = \mathbb{P}(A|B) \times \mathbb{P}(B) = \mathbb{P}(A) \times \mathbb{P}(B)$$

Conclusión: A y B son independientes ssi $\mathbb{P}(A|B) = \mathbb{P}(A)$

Regla de la cadena

Considere una colección de n eventos $\{A_1, A_2, \ldots, A_n\}$

$$\begin{aligned}
\mathbb{P}(A_1, A_2) &= \mathbb{P}(A_1|A_2) \times \mathbb{P}(A_2) \\
\mathbb{P}(A_1, A_2, A_3) &= \mathbb{P}(A_1|A_2, A_3) \times \mathbb{P}(A_2, A_3) \\
&= \mathbb{P}(A_1|A_2, A_3) \times \mathbb{P}(A_2|A_3) \times \mathbb{P}(A_3) \\
\mathbb{P}(A_1, A_2, A_3, A_4) &= \mathbb{P}(A_1|A_2, A_3, A_4) \times \mathbb{P}(A_2, A_3, A_4) \\
&= \mathbb{P}(A_1|A_2, A_3, A_4) \times \mathbb{P}(A_2|A_3, A_4) \times \mathbb{P}(A_3, A_4) \\
&= \mathbb{P}(A_1|A_2, A_3, A_4) \times \mathbb{P}(A_2|A_3, A_4) \times \mathbb{P}(A_3|A_4) \times \mathbb{P}(A_4) \\
\mathbb{P}(A_1, \ldots, A_k) &= \Pi_{i=1}^{k} \mathbb{P}(A_i|A_{i+1}, \ldots, A_k)
\end{aligned}$$

Variables aleatorias

Considere la siguiente propuesta:

Usted paga n bolívares para tener el chance de lanzar una moneda. Si la moneda sale cara, usted recibe un pago m_H y si la moneda sale cruz usted recibe m_T (ambos pagos en bolívares)

¿Vale la pena aceptar la propuesta?

Utilizaremos variables aleatorias y cálculo de esperanza para responder la pregunta

Variables aleatorias

Una variable aleatoria X sobre un espacio $(\Omega, \mathcal{F}, \mathbb{P})$ es una **función** desde el espacio muestral a los reales (i.e. $X : \Omega \to \mathbb{R}$)

Definimos,

$$\mathbb{P}(X = x) = \mathbb{P}(\{w \in \Omega : X(w) = x\})$$

En el ejemplo, $\Omega = \{H, T\}$ y definimos la variable aleatoria $Z : \Omega \to \mathbb{R}$ igual al **beneficio neto** obtenido de aceptar la propuesta:

$$\begin{aligned}
Z(H) &= m_H - n \\
Z(T) &= m_T - n
\end{aligned}$$

E.g., para $m_H \neq m_T$, $\mathbb{P}(Z = m_H - n) = \mathbb{P}(\{H\})$

Esperanza

Dado $(\Omega, \mathcal{F}, \mathbb{P})$ y variable aleatoria $X : \Omega \to \mathbb{R}$, la **esperanza** de X se define como:

$$\mathbb{E}[X] \;=\; \textstyle\sum_{\omega\in\Omega} X(\omega)\,\mathbb{P}(\{\omega\}) \;=\; \sum_x x\,\mathbb{P}(X=x)$$

La esperanza cuantifica el valor esperado (promedio) de la variable aleatoria X:

Si realizamos el experimento muchas veces, y promediamos los valores de X obtenidos, el promedio tiende a $\mathbb{E}[X]$ a medida que el número de repeticiones del experimento se incrementa

Resolución del problema de apuesta

El beneficio neto Z de aceptar la propuesta es:

$$\begin{aligned} Z(H) &= m_H - n \\ Z(T) &= m_T - n \end{aligned}$$

El beneficio neto esperado es:

$$\begin{aligned} \mathbb{E}[Z] &= \textstyle\sum_{\omega\in\Omega} Z(\omega)\,\mathbb{P}(\{\omega\}) \\ &= Z(H)\,\mathbb{P}(\{H\}) + Z(T)\,\mathbb{P}(\{T\}) \\ &= (m_H - n) \times \tfrac{1}{2} + (m_T - n) \times \tfrac{1}{2} \\ &= \tfrac{1}{2}(m_H + m_T) - n \end{aligned}$$

Entonces, $\mathbb{E}[Z] > 0$ cuando $m_H + m_T > 2n$

Independencia de variables aleatorias

Dos variables aleatorias X, Y definidas sobre el mismo espacio $(\Omega, \mathcal{F}, \mathbb{P})$ son **independientes** ssi

$$\mathbb{P}(X=x, Y=y) \;=\; \mathbb{P}(X=x) \times \mathbb{P}(Y=y)$$

para todo $x, y \in \mathbb{R}$

Una colección $\{X_i\}_{i=1}^n$ de variables aleatorias definidas sobre el mismo espacio es independiente ssi

$$\mathbb{P}(\cap_{i=1}^n X_i = x_i) \;=\; \textstyle\prod_{i=1}^n \mathbb{P}(X_i = x_i)$$

para todo $x_1, x_2, \ldots, x_n \in \mathbb{R}$

Propiedades de esperanza

- $\mathbb{E}[\alpha X + \beta Y] \;=\; \alpha\mathbb{E}[X] + \beta\mathbb{E}[Y]$

- Si $Y = g(X)$ para $g : \mathbb{R} \to \mathbb{R}$, entonces Y es variable aleatoria y

$$\mathbb{E}[Y] = \textstyle\sum_y y\mathbb{P}(Y=y) = \sum_x g(x)\mathbb{P}(X=x)$$

- Si X e Y son independientes, $\mathbb{E}[XY] \;=\; \mathbb{E}[X] \times \mathbb{E}[Y]$

Variables aleatorias indicadoras

Una variable aleatoria $X : \Omega \to \mathbb{R}$ tal que $Rg(X) = \{0, 1\}$ se llama **variable aleatoria indicadora**

Calculemos su esperanza:

$$
\begin{aligned}
\mathbb{E}[X] &= \textstyle\sum_{\omega\in\Omega} X(\omega)\,\mathbb{P}(\{\omega\}) \\
&= \textstyle\sum_{\omega\in\Omega, X(\omega)=1} \mathbb{P}(\{\omega\}) \\
&= \mathbb{P}(\{\omega \in \Omega : X(\omega) = 1\}) \\
&= \mathbb{P}(E)
\end{aligned}
$$

donde E es el evento $E = \{\omega \in \Omega : X(\omega) = 1\} = \{X = 1\}$

Cota para la probabilidad de la unión

Considere una colección de eventos $\{E_1, E_2, \ldots, E_n\}$

Si los eventos son disjuntos dos a dos,

$$\mathbb{P}(\cup_{1\leq i\leq n} E_i) = \sum_{1\leq i\leq n} \mathbb{P}(E_i)$$

En cualquier caso,

$$\mathbb{P}(\cup_{1\leq i\leq n} E_i) \leq \sum_{1\leq i\leq n} \mathbb{P}(E_i)$$

Esperanza de variables enteras

Considere una v.a. entera no-negativa X; i.e. con rango $\{0, 1, 2, \ldots\}$

$$\mathbb{E}[X] = \textstyle\sum_{x\geq 0} x\,\mathbb{P}(X = x) = \sum_{x\geq 1} x\,\mathbb{P}(X = x)$$

$$
\begin{array}{lllll}
\mathbb{P}(X=1) & & & & \\
\mathbb{P}(X=2) & \mathbb{P}(X=2) & & & \\
\mathbb{P}(X=3) & \mathbb{P}(X=3) & \mathbb{P}(X=3) & & \\
\mathbb{P}(X=4) & \mathbb{P}(X=4) & \mathbb{P}(X=4) & \mathbb{P}(X=4) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\hline
\mathbb{P}(X\geq 1) & \mathbb{P}(X\geq 2) & \mathbb{P}(X\geq 3) & \mathbb{P}(X\geq 4) & \cdots
\end{array}
$$

$$\mathbb{E}[X] = \textstyle\sum_{x\geq 1} \mathbb{P}(X \geq x)$$

Regla de probabilidad total

Consider un evento A y una colección de eventos $\{B_i\}_i$ que **particionan el espacio muestral** Ω

$$\mathbb{P}(A) = \sum_i \mathbb{P}(A \cap B_i)$$

La prueba es sencilla ya que la colección de eventos $\{A \cap B_i\}_i$ es una colección disjunta dos a dos y $\cup_i (A \cap B_i) = A$

Desigualdad de Markov

La desigualdad de Markov permite acotar la probabilidad de una variable X no negativa utilizando la esperanza de X

Para todo $t > 0$,

$$P(X \geq t) \;\leq\; \frac{\mathbb{E}[X]}{t}$$

Ejemplos de análisis probabilístico

Paradoja del día de nacimiento 1/3

¿Cuántos estudiantes deben haber en un salón para que la probabilidad que dos de ellos nazcan el mismo día sea $> \frac{1}{2}$?

Asumiremos que el año tiene $n = 365$ días (obviando años bisiestos) e indizamos a las personas con los enteros $1, 2, \ldots, k$

Sea $b_i \in \{1, 2, \ldots, n\}$ el día de nacimiento de la persona i. Asumimos $\mathbb{P}(b_i = r) = 1/n$ y $\mathbb{P}(b_i = r \;\&\; b_j = s) = 1/n^2$ para todo i, j, r, s

La probabilidad que las personas i y j nazcan el mismo día es

$$\mathbb{P}(b_i = b_j) \;=\; \sum_{d=1}^{n} \mathbb{P}(b_i = d \;\&\; b_j = d) \;=\; \sum_{d=1}^{n} \frac{1}{n^2} \;=\; \frac{1}{n}$$

Paradoja del día de nacimiento 2/3

Ahora acotaremos $\mathbb{P}(B_m)$ donde B_m es el evento que los **días de nacimiento de las primeras m personas son todos distintos**

Observe $B_m = B_{m-1} \cap A_m$ donde A_m es el evento que las personas $1, \ldots, m-1$ nacen en días distinto a b_m. Entonces, $B_m = \bigcap_{i=1}^{m} A_i$

Para responder la pregunta queremos calcular el número k de personas tal que $1 - \mathbb{P}(B_k) \geq \frac{1}{2}$. Es decir, k tal que $\mathbb{P}(B_k) \leq \frac{1}{2}$

$$\begin{aligned}
\mathbb{P}(B_k) &= \mathbb{P}(B_1) \times \mathbb{P}(A_2|B_1) \times \mathbb{P}(A_3|B_2) \times \cdots \times \mathbb{P}(A_k|B_{k-1}) \\
&= 1 \times \tfrac{n-1}{n} \times \tfrac{n-2}{n} \times \cdots \times \tfrac{n-k+1}{n} \\
&= 1 \times (1 - \tfrac{1}{n}) \times (1 - \tfrac{2}{n}) \times \cdots \times (1 - \tfrac{k-1}{n}) \\
&\leq e^{-1/n} \times e^{-2/n} \times \cdots \times e^{-(k-1)/n} \qquad [\alpha < 1 \implies 1 - \alpha \leq e^{-\alpha}] \\
&= e^{-\frac{1}{n}\sum_{i=1}^{k-1} i} \;=\; e^{-k(k-1)/2n} \;\leq\; \tfrac{1}{2}
\end{aligned}$$

cuando $-k(k-1)/2n \leq \ln(1/2)$. Para $n = 365$, basta $k \geq 23$

Paradoja del día de nacimiento 3/3

Resolvamos ahora usando **variables indicadoras y la linealidad de la esperanza**. Considere la v.a. indicadora $X_{ij} = \mathbb{I}\{b_i = b_j\}$:

$$\mathbb{E}[X_{ij}] \;=\; \mathbb{P}(X_{ij} = 1) \;=\; \mathbb{P}(b_i = b_j) \;=\; 1/n$$

Sea X el número de pares de personas distintas con el mismo día de nacimiento: i.e. $X = \sum_{1 \leq i < j \leq k} X_{ij}$

$$\mathbb{E}[X] \;=\; \textstyle\sum_{1\leq i<j\leq k} \mathbb{E}[X_{ij}] \;=\; \sum_{1\leq i<j\leq k} \frac{1}{n} \;=\; \binom{k}{2}\frac{1}{n} \;=\; \frac{k(k-1)}{2n}$$

Si $\mathbb{E}[X] \geq 1$ entonces existe al menos un i y j tal que $X_{ij} = 1$

$\mathbb{E}[X] \geq 1$ cuando $k(k-1)/2n \geq 1$ lo que implica $k \geq \frac{1}{2}(1 + \sqrt{1+8n})$

Para $n = 365$, cuando $k \geq 28$ esperamos que exista al menos 1 par de personas que cumplan el mismo día

Más fácil pero la cota $k \geq 28$ es menos ajustada que $k \geq 23$

Pelotas y cajas

Considere el proceso de lanzar n **pelotas idénticas** en b **cajas distintas**, numeradas $1, 2, \ldots, b$

Los **lanzamientos son independientes** y la probabilidad de que una pelota caiga en una caja cualquiera es $1/b$

– ¿Cuántas pelotas caen (en promedio) en una caja dada?

Respuesta: n/b pelotas por caja en promedio

– ¿Cuántas pelotas debemos lanzar (i.e. determinar n) hasta que 1 pelota caiga en una caja dada?

Respuesta: debemos lanzar b pelotas en promedio

– ¿Cuántas pelotas debemos lanzar (i.e. determinar n) hasta que toda caja contenga al menos 1 pelota? **(Coleccionista de cupones)**

Respuesta: debemos lanzar $b \ln b$ pelotas en promedio

Coleccionista de cupones 1/2

Cada vez que una pelota cae en una caja vacía lo llamamos un **hit**

Queremos calcular el número esperado de lanzamientos hasta obtener b hits

Particionamos los lanzamientos en episodios:

– el primer episodio consiste en los lanzamientos hasta el primer hit (incluyendo el hit)

– el i-ésimo episodio consiste en los lanzamientos después de $(i-1)$-ésimo hit y hasta el i-ésimo hit (incluyendo el hit)

Sea n_i el número de lanzamientos en el i-ésimo episodio. El número total de lanzamientos para obtener b hits es $\sum_{i=1}^{b} n_i$

$$\mathbb{E}\left[\textstyle\sum_{i=1}^{b} n_i\right] \;=\; \textstyle\sum_{i=1}^{b} \mathbb{E}[n_i] \;=\; \sum_{i=1}^{b} \frac{b}{b-i+1} \;=\; b\sum_{i=1}^{b}\frac{1}{i} \;=\; b(\ln b + O(1))$$

donde $\mathbb{E}[n_i] = b/(b-i+1)$ (ver próximo slide)

Coleccionista de cupones 2/2

Distribución y esperanza de la variable n_i

Sea p_i la probabilidad de que un lanzamiento en el i-ésimo episodio sea un hit y $q_i = 1 - p_i$

$$\mathbb{P}(n_i \geq k) \;=\; \mathbb{P}(n_i > k-1) \;=\; q_i^{k-1}$$

$$\mathbb{E}[n_i] \;=\; \sum_{k\geq 1} \mathbb{P}(n_i \geq k) \;=\; \sum_{k\geq 1} q_i^{k-1} \;=\; \sum_{k\geq 0} q_i^{k} \;=\; \frac{1}{1-q_i} \;=\; \frac{1}{p_i}$$

Observe que $p_i = (b-i+1)/b$ ya que durante el i-ésimo episodio, $i-1$ cajas tiene pelotas y $b-i+1$ cajas no tienen pelotas

Resumen

- Elementos de teoría de probabilidades:
  - espacio de probabilidades: espacio muestral, eventos, función de probabilidad
  - independencia de eventos, probabilidad condicional y regla de la cadena
  - variables aleatorias e independencia de variables
  - esperanza y propiedades
  - cota para la unión, regla de probabilidad total y desigualdad de Markov
- Ejemplos de análisis probabilístico

Ejercicios (1 de 2)

1. (8.4-3) Sea X una variable aleatoria que es igual al número de caras en dos lanzadas de una moneda insesgada. Calcule $\mathbb{E}[X^2]$ y $\mathbb{E}[X]^2$
2. (5.2-3) Use variables indicadoras para computar el valor esperado de la suma de n dados lanzados (insesgados)
3. (5.2-4) Use variables indicadoras para resolver lo siguiente. Cada uno de los n clientes en un restaurant dejan sus sombreros con el guardaropas. Al salir, el guardaropas devuelve los sombreros totalmente al azar. ¿Cuál es el número esperado de clientes que obtienen su sombrero al salir del restaurant?

Ejercicios (2 de 2)

4. (5.2-5) Sea $A[1 \dots n]$ un arreglo con n enteros distintos. Si $i < j$ y $A[i] > A[j]$, decimos que el par (i, j) es una inversión de A. Suponga que los elementos de A forman una permutación uniforme sobre $\{1, 2, \dots, n\}$. Use variables indicadoras para calcular el número esperado de inversiones en A
5. Responda la primera y segunda pregunta para el problema de las pelotas y las cajas