Tarea 3 - Análisis de datos
Edgar Osvaldo López Zúñiga
Giovanni Gamaliel López Padilla

Problema 10

Considera el siguiente grafo:

Figura 1: Mapa del sitio con los nodos nombrados

Cada calle (arista) entre dos nodos esté bloqueda por una manifestación con probabilidad p. Supongamos que todos son eventos independientes.

Problema 10a

Calcula la probalidad de poder caminar desde la puerta trasera del Teatro Juarez (nodo más abajo) al café conquistador (nodo más hacia arriba).
Llamaremos $P(X_1 \rightarrow X_2)$ a la probabilidad de poder caminar desde un nodo X_1 hasta un nodo X_2 sin repetir nodos intermedios y tomando en cuenta todos los caminos posibles entre X_1 y X_2.
En la figura 1 se muestran los nodos, a los que hemos nombrado $A, B, ..., J$ para mayor simplicidad. Además, llamaremos $p_{\alpha\beta}$ a la probabilidad de que el vértice entre los nodos α y β esté libre, es decir que se pueda caminar por la calle que une los dos puntos.
A nosotros nos interesa la probabilida de que haya camino libre desde el nodo de más abajo (el nodo A) hasta el nodo de más arriba (nodo J), es decir, nos interesa $P(A \rightarrow J)$. Si tomamos en cuenta que todos los caminos desde A hasta J, pasan por D podemos afirmar lo siguiente:

$$P(A \rightarrow J) = P(D \rightarrow J)P(A \rightarrow D)$$

Es decir, para que se pueda caminar desde A hasta J, tiene que haber camino libre entre A y D y entre D y J. De la misma manera podemos decir que para que los nodos D y J estén conectados, debe haber paso entre D y H y entre H y J, o podemos caminar desde D a I y después de I a J. Esto, lo podemos escribir de la siguiente manera:

$$P(A \rightarrow J) = [p_{HJ}P(D \rightarrow H) + p_{IJ}P(D \rightarrow I)]\, P(A \rightarrow D)$$

A partir de ahora procederemos a simplificar esta expresión utilizando la misma lógica que en los pasos anteriores:
Si necesitamos que los caminos de X_i a X_j y los caminos de X_j a X_k estén desbloqueados, la probabilidad de este evento es el producto de las probabilidades. Y si necesitamos que haya paso de X_i a X_k o de X_k a X_k, la probabilidad de este evento, es la suma de las probabilidades.

$$
\begin{aligned}
P(A \to J) &= [p_{HJ}P(D \to H) + p_{IJ}P(D \to I)]\, P(A \to D) \\
&= [p_{HJ}p_{FH}P(D \to F) + p_{IJ}P(D \to I)]\, P(A \to D) \\
&= [p_{HJ}p_{FH}\,(p_{DF} + p_{GF}p_{EG}p_{DE} + p_{IF}p_{GI}p_{EG}p_{DE}) + p_{IJ}\,(p_{FI}P(D \to F) + p_{GI}P(D \to G))]\, P(A \\
&= [p_{HJ}p_{FH}\,(p_{DF} + p_{GF}p_{EG}p_{DE} + p_{IF}p_{GI}p_{EG}p_{DE}) \\
&\quad + p_{IJ}\,(p_{FI}(p_{DF} + p_{GF}p_{EG}p_{DE}) + p_{GI}(p_{FG}p_{DF} + p_{EG}p_{DE}))] P(A \to D)
\end{aligned}
$$

Ahora, notemos que la probabilidad de que una calle está bloqueada es la misma en todos los casos, por lo que podemos simplificar todos los $p_{\alpha\beta}$ haciendo lo siguiente:

$$q = p_{\alpha\beta} = 1 - p \ \forall \alpha, \beta \in \{A,\ B,\ ...,\ J\}$$

Así, expresión de la probabilidad de que haya camino libre entre A y J se convierte en lo siguiente:

$$
\begin{aligned}
&P(A \to J) = \left[q^2(q + q^3 + q^4) + q(q(q + q^3) + q(q^2 + q^2))\right] P(A \to D), \\
&\text{además } P(A \to D) = [p_{AB}p_{BD} + p_{AC}p_{CD}] = 2q^2 \\
&\text{lo que nos lleva a} \\
&P(A \to J) = 2q^2\left[q^3 + q^5 + q^6 + q^3 + q^5 + 2q^4\right] \\
&\qquad\qquad = 2q^8 + 4q^7 + 4q^6 + 4q^5
\end{aligned}
$$

Por lo tanto, la probabilidad de poder caminar desde la puerta trasera del Teatro Juarez al café conquistador es:

$$P(A \to J) = 2(1-p)^8 + 4(1-p)^7 + 4(1-p)^6 + 4(1-p)^5$$

Problema 10b

Decimos que un nodo del grafo está aislado si todas las aristas que llegan (o salen) están bloqueadas. Calcula el promedio del número de nodos aislados.
Sea i el número de aristas que salen del nodo y a_i el número de nodos que tienen i aristas. El promedio de nodos aislados puede ser calculado como la suma del promedio de los a_i nodos, entonces:

$$
\begin{aligned}
E(A) &= \sum_i E(A_i) \\
E(A) &= \sum_i a_i P(a_i) \qquad (1)
\end{aligned}
$$

donde $P(a_i)$ es la probabilidad que todas las aristas del nodo a_i esten bloqueadas. Como la probabilidad que una arista este bloqueda es independiente de las demas, entonces se obtiene que $P(a_i)$ es:

$$
\begin{aligned}
P(a_i) &= \prod_i P(a_{ii}) \\
&= P(a_{i1})P(a_{i2})\ldots P(a_{ii})
\end{aligned}
$$

Como la probabilidad es la misma en todas las aristas entonces:

$$P(a_i) = q^i \tag{2}$$

Entonces, usando la ecuación 2 en la ecuación 1 se obtiene que:

$$\begin{aligned} E(A) &= \sum_i a_i P(a_i) \\ &= \sum_i a_i q^i \end{aligned} \tag{3}$$

Los nodos que componen a los conjuntos a_i del mapa (figura 1) son los siguientes:

$$a_i = \begin{cases} a_2 = & \Big\{A, B, C, H, J, E\Big\} \\ a_3 = & \Big\{I, G\Big\} \\ a_4 = & \Big\{D, F\Big\} \end{cases}$$

Por lo tanto, los coeficientes son 6, 2, 2 para los nodos a_2, a_3, a_4 respectivamente. Introduciendo esta información en la ecuación 3 se obtiene:

$$E(A) = 6q^2 + 2q^3 + 2q^4 \tag{4}$$

lo cual es el promedio que un nodo este aislado.