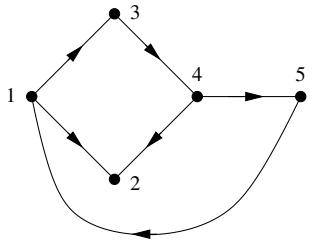
- Contrôle continu de Mathématiques du Web -

- Durée : 1 heure 30 -

Aucun document autorisé, barême indicatif

- Exercice 1 - HITS - *8 points*
Au cours de cet exercice, on considère la portion suivante du graphe du web. On note D ce graphe orienté.



Une recherche utilisant l'algorithme de Hits est lancée sur D.

1. Si A désigne la matrice d'adjacence du graphe D, on rappelle que la matrice 'authority' de D est la matrice ${}^tA.A$. Donner l'interprétation graphique (*i.e.* sur le graphe) des cœfficients de cette matrice.
2. Calculer la matrice 'authority' de D.
3. Calculer le score et le classement 'authority' pour la requête effectuée (on donnera le score sous forme d'un vecteur x, pas forcément normé, dont les coordonnées sont les scores des pages impliquées).
4. Pour une matrice carrée B quelconque, montrer que si x est vecteur propre de ${}^tB.B$ pour la valeur propre λ, alors Bx est vecteur propre de $B.{}^tB$ pour la même valeur propre λ.
5. En déduire, en un seul calcul vectoriel, le score et le classement 'hub' de la requête.

- Exercice 2 - *Question de cours/tp - 4 points*
On considère une variable aléatoire entière X dont la loi de probabilités est donnée par la suite $(u_k)_{k\in\mathbb{N}}$ (*i.e.* pour tout $k \in \mathbb{N}$, on a $P(X = k) = u_k$).
Écrire en pseudo-code (type C) une procédure permettant de simuler le tirage aléatoire d'une valeur suivant la loi de X. On pourra faire appel à une fonction `rand()` renvoyant un réel choisi uniformément entre 0 et 1.

- Exercice 3 - Modèle aléatoire de graphe - *8 points*
On considère le modèle aléatoire suivant : pour construire le graphe G sur les sommets $\{1, \dots n\}$, on ajoute les arêtes $\{\{i, i+1\} : i = 1, \dots, n-1\} \cup \{1, n\}$ pour obtenir un cycle, puis pour chaque sommet i on ajoute l'arête $\{i, x_i\}$ où x_i est choisi uniformément dans $\{1, \dots, n\} \setminus i$. On autorise les arêtes doubles et triples mais on n'en tiendra pas compte dans les calculs.

1. Justifier rapidement que pour tout i, on ait $d(i) \geq 3$, où $d(i)$ désigne le degré de i dans le graphe G, et que pour tout $j \neq i$, on ait $P(x_j = i) = \frac{1}{n-1}$.

2. En déduire, la loi de degré du sommet i, c'est-à-dire, calculer pour tout $k \geq 0$ la valeur de $P(d(i) = k + 3)$.
3. De quelle loi classique cette distribution de degré se rapproche-t-elle ? Rapidement, à votre avis, que dire quand n devient très grand ?
4. Donner le degré moyen dans le graphe G.
5. Dans cette question, on veut montrer que le diamètre du graphe est inférieur à $\frac{3}{2}\sqrt{n}$ en moyenne. Pour deux sommets i et j du graphe G, on note $d(i, j)$ la distance de i à j, c'est-à-dire la longueur d'un plus court chemin de i à j.
   (a) Soit i un sommet, on définit $B(i) = \{j \ : \ d(i, j) \leq \frac{1}{2}\sqrt{n}\}$. Justifier que $|B(i)| \geq \sqrt{n}$. Pour un sommet $j \notin B(i)$, calculer la probabilité que $x_j \in B(i)$, en fonction de $|B(i)|$.
   (b) Soit k un sommet hors de $B(i)$, et $P = p_1, \ldots, p_q$ un chemin de $p_1 = i$ à $p_q = k$. On suppose que $p_j \in B(i)$ pour $j = 1, \ldots, l$ et que $p_j \notin B(i)$ pour $j = l + 1, \ldots, q$. Calculer le nombre moyen de sommet p_j avec $j > l + 1$ ayant un voisin dans $B(i)$.
   (c) *Hors-barême* Conclure, qu'en moyenne, la distance entre i et k est inférieure ou égale à $\frac{3}{2}\sqrt{n}$.