** Please hand in solutions to questions 1 and 3 on this sheet. This counts towards your final mark for Level M students **

1. Let $p$ and $q$ be discrete probability distributions on a finite set $\Omega = \{1, 2, \ldots, n\}$. Recall that the total variation distance between $p$ and $q$ is defined as

$$d_{TV}(p, q) = \max_{S \subseteq \Omega} |p(S) - q(S)|,$$

where $p(S) = \sum_{i \in S} p_i$ and $q(S)$ is defined similarly.

(a) Show that $d_{TV}(p, q) = \frac{1}{2} \|p - q\|_1$, where $\|p - q\|_1$ is defined as $\sum_{i=1}^n |p_i - q_i|$. (Hint. Show that the maximum in the definition of total variation distance is attained by the set $S = \{i : p_i \geq q_i\}$.)

(b) Using the answer to the last part or otherwise, show that $\|p - q\|_1 \leq 2$ for all probability distributions $p$ and $q$. Give an example where equality holds.

(c) Show that $\|p - q\|_1 \leq \sqrt{n} \|p - q\|_2$, where $\|x\|_2$ is defined as $\sqrt{x_1^2 + \ldots + x_n^2}$. (Hint. Use the Cauchy-Schwarz inequality.)

(d) Show that $\|p - q\|_2 \leq \sqrt{2}$ for any probability distributions $p$ and $q$. Give an example where equality holds.

2. Compute the total variation distance between the two distributions in each of the following examples:

(a) Binomial(2, $\frac{1}{2}$) and uniform on $\{0, 1, 2\}$.

(b) Binomial(2, $\frac{1}{2}$) and Poisson(1).

(c) Exponential(1) and Uniform[0,1].

(d) Exponential(1) and Exponential(2).

3. Let $S_n$ be the star graph on $n$ nodes consisting of a single hub node connected to each of $n - 1$ leaves; there are no edges between leaves. Consider the continuous time random walk on this graph generated as follows: there are $n - 1$ independent Poisson processes, $\{N_e(t), t \geq 0\}$, one on each edge of the graph. Each of these Poisson processes has rate 1. If the Poisson process $N_e(\cdot)$ has an increment at time $t$ and the walker is at the vertex on one end of this edge, it moves to the vertex at the other end. An equivalent description is that, if the walker is at a leaf, it moves to the hub at rate 1; if it is at the hub, it moves at rate $n - 1$, choosing a leaf uniformly at random to move to.
(a) The position $X_t$ of the random walk evolves as a continuous time Markov chain on the set of vertices. Write down the transition rate matrix of this Markov chain. How is it related to the Laplacian matrix of the graph $S_n$?

(b) Show that the uniform distribution on all nodes is an invariant distribution for the Markov chain in part (a), and hence that it is the unique invariant distribution.

(c) The conductance of a graph $G$ is defined as

$$
\Phi(G) = \min_{S \subset V: S \neq \emptyset} \frac{|E(S, S^c)|}{\frac{1}{n} |S| \cdot |S^c|},
$$

where $E(S, S^c)$ denotes the set of all edges consisting of one vertex in $S$ and the other in its complement $S^c$, and the minimum is taken over all subsets of the vertex set $V$ other than the empty set and the set of all vertices.

Compute the conductance of the star graph.

(d) Obtain a lower bound on $\lambda_2$, the second smallest eigenvalue of the Laplacian of $S_n$, using Cheeger’s inequality and the answer to part (c).

(e) It is known that the total variation distance between the distribution of the random walk position at time $t$, which we denote $p(t)$, and the invariant distribution $\pi$ is bounded as follows:

$$
\|p(t) - \pi\|_{TV} := \frac{1}{2} \sum_{i=1}^{n} |p_i(t) - \pi_i| \leq \sqrt{n} e^{-\lambda_2 t}.
$$

Let $\epsilon > 0$ be a given constant. Using the lower bound on $\lambda_2$ computed in part (d), find the smallest value of $t$ for which you can guarantee that $\|p(t) - \pi\|_{TV} \leq \epsilon$.

4. Consider a directed graph with adjacency matrix

$$
A = \begin{pmatrix}
0 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0
\end{pmatrix}
$$

Assuming that this represents a small web graph, compute the page rank of each of the nodes.