We will denote the set $\{1, 2, \ldots n\}$ as $[n]$.

Question 1 *Prove that* $\max_{A \subseteq [n]} |p(A) - q(A)| = \frac{\sum_{i=1}^{n} |p(i) - q(i)|}{2}$

$$|P(A) - Q(A)| \equiv \left| \sum_{a \in A} p(a) - q(a) \right|$$

Note that to maximise the above quantity, we can choose to maximise either positive values or negative values, since it is surrounded by $| \, |$. Let us arbitrarily choose to maximise positive values (the solution is symmetric).

In that case, we need to ensure that we pick $a_0 \in [n]$ such that $[P(a_0) - Q(a_0) > 0]$. This forces us to pick $A \equiv \{a \in [n] \ : \ P(a) - Q(a) > 0\}$. Now, define $\bar{A} \equiv \{a \in [n] \ : \ P(a) - Q(a) \leqslant 0\}$. That is, $A \cap \bar{A} = \emptyset$, $A \cup \bar{A} = [n]$.

Recall that $\sum_a p(a) = 1$, $\sum_a q(a) = 1$.

$$\begin{aligned}
&\sum_{i=1}^{n} |p(i) - q(i)| \\
&= \sum_{a \in A} |p(a) - q(a)| + \sum_{\bar{a} \in \bar{A}} |p(\bar{a}) - q(\bar{a})| \\
&= \sum_{a \in A} (p(a) - q(a)) + \sum_{\bar{a} \in \bar{A}} (q(\bar{a}) - p(\bar{a}))
\end{aligned}$$

Question 2 *Prove that on a finite DAG, at least one vertex has no incoming edges.*

Proof sketch: Build a chain by picking a vertex v_0. If this vertex has no incoming edges, then we are done. If not, pick a predecessor v_1 such that $v_1 \to v_0$. Now, attempt to pick a predecessor of v_1. If it has no predecessor, we are done. If not, we must have a new vertex v_2. The crucial point is that this v_2 is *not equal to* v_0, v_1. For if it were, we would have a cycle of the form $(v_2 \to v_1 \to v_0 \to v_2)$. or of the form $(v_2 \to v_1 \to v_2)$.

Hence, we have a *decreasing measure*, the number of available vertices to extend the chain is the number of vertices in the graph minus the length of the chain. This is a finite number, and cannot be less than 0. Thus this process must terminate, yielding a final vertex at some point that has no predecessor.

Question 3 *Consider the three variable distribution* $P(a, b, c) = P(a|b)P(b|c)P(c)$ *where all variables are binary. how many parameters are needed to specify a distribution of this form?*

For each $c = 0, c = 1$, we need the values of $P(b = 1|c = 0), P(b = 1|c = 1)$ which is 2 parameters. The other two $P(b = 0|c = 0), P(b = 0|c = 1)$ can be calculated.

For each $b = 0, b = 1$, we need the values of $P(a = 1|b = 0), P(a = 1|b = 1)$ which is 2 parameters.

So, in total, we need $2 + 2 + 1 = 5$ parameters.