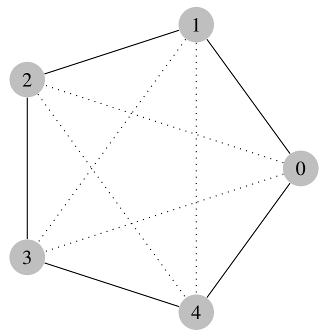
1 Clarification on Expected value

Consider a connected graph of 5 nodes (we assume that you can travel to any node from any other node). We can assume without loss of generality that the nodes are numbered so that the solution to the TSP is the tour $\tau = 012340$ and for sake of clarity let the nodes be distributed evenly around a unit circle. We represent the solution to the TSP problem with solid lines and all other possible paths with dotted lines as below:



Thinking of this in terms of each path though (letting the probability of a directed path from nodes a to node b being realized be given as $P(\overline{ab})$...

- $P(\overline{01}) = p_0 p_1$
  - This is the path from node 0 to node 1. Since this is in the solution to the TSP, it will occur every time nodes 0 and 1 are realized.
- $P(\overline{02}) = p_0 p_2 (1 - p_1)$
  - This is the path from node 0 to node 2. This is not in the solution to the TSP, and so will only occur when nodes 0 and 2 are realized, but node 1 is not realized.
- $P(\overline{03}) = p_0 p_3 (1 - p_1)(1 - p_2)$
  - This is the path from node 0 to node 3. This is not in the solution to the TSP, and so will only occur when nodes 0 and 3 are realized, but nodes 1 and 2 are not realized.
- $P(\overline{04}) = p_0 p_4 (1 - p_1)(1 - p_2)(1 - p_3)$
  - This is the path from node 0 to node 4. This is not in the solution to the TSP, and so will only occur when nodes 0 and 4 are realized, but nodes 1, 2, and 3 are not realized.
- $P(\overline{12}) = p_1 p_2$
  - This is the path from node 1 to node 2. Since this is in the solution to the TSP, it will occur every time nodes 1 and 2 are realized.
- $P(\overline{13}) = p_1 p_3 (1 - p_2)$
  - This is the path from node 1 to node 3. This is not in the solution to the TSP, and so will only occur when nodes 1 and 3 are realized, but node 2 is not realized.
- $P(\overline{14}) = p_1 p_4 (1 - p_2)(1 - p_3)$
  - This is the path from node 1 to node 4. This is not in the solution to the TSP, and so will only occur when nodes 1 and 4 are realized, but nodes 2 and 3 are not realized.
- $P(\overline{23}) = p_2 p_3$
  - This is the path from node 2 to node 3. Since this is in the solution to the TSP, it will occur every time nodes 2 and 3 are realized.
- $P(\overline{24}) = p_2 p_4 (1 - p_3)$

- This is the path from node 2 to node 4. This is not in the solution to the TSP, and so will only occur when nodes 2 and 4 are realized, but node 3 is not realized.

- $P\left(\overline{34}\right) = p_3 p_4$
  - This is the path from node 3 to node 4. Since this is in the solution to the TSP, it will occur every time nodes 3 and 4 are realized.
- $P\left(\overline{10}\right) = p_1 p_0 (1 - p_2)(1 - p_3)(1 - p_4)$
  - This is the trip returning to node 0 from node 1. Since this is not in the solution to the TSP, it will only occur if nodes 0 and 1 are realized but nodes 2, 3, and 4 are not realized.
- $P\left(\overline{20}\right) = p_2 p_0 (1 - p_3)(1 - p_4)$
  - This is the trip returning to node 0 from node 2. Since this is not in the solution to the TSP, it will only occur if nodes 0 and 2 are realized but nodes 3 and 4 are not realized.
- $P\left(\overline{30}\right) = p_3 p_0 (1 - p_4)$
  - This is the trip returning to node 0 from node 3. Since this is not in the solution to the TSP, it will only occur if nodes 0 and 3 are realized but node 4 are not realized.
- $P\left(\overline{40}\right) = p_4 p_0$
  - This is the trip returning to node 0 from node 4. Since this is in the solution to the TSP, it will occur when nodes 0 and 4 are realized.

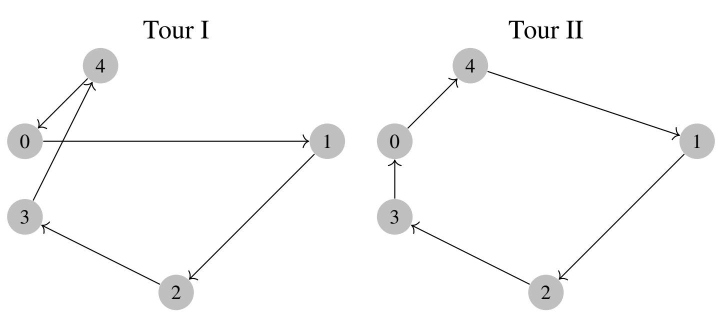
Letting d_{ab} be the distance between nodes a, b, we get that the expected value of any path between nodes a, b is given by $P\left(\overline{ab}\right) \cdot d_{ab}$ and so the expected value of of a tour is given by $\sum P\left(\overline{ab}\right) \cdot d_{ab}$ for all nodes a, b in the graph.

More concisely, we have that the formula for the expected value of this tour, given probabilities $P = \left[p_0 = 1, p_1, p_2, p_3, p_4\right]$ (where $p_0 = 1$ due to the supply node always occuring) is given by

$$E(L_\tau) = \sum_{j=1}^{n} p_j d_{0,j} \prod_{k=1}^{j-1} (1 - p_k) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} p_i p_j d_{i,j} \prod_{k=i+1}^{j-1} (1 - p_k) + \sum_{i=1}^{n} p_i d_{i,0} \prod_{k=i+1}^{n} (1 - p_k) \tag{1}$$

2 Illustrative Example Explained

Going back to the illustrative example from Campbell/Thomas,



We have nodes at $\{(0, 0), (4, 0), (2, -2), (0, -1), (1, 1)\}$ with the probabilities of the nodes occurring, $P = \left[p_0 = 1, 0.1, 1, 1, 0.5\right]$.

Letting the cost for traveling between two nodes be equal to the distance between them, we obtain the following cost/distance matrix:

$$d_{ij} = \begin{pmatrix} 0 & 4 & 2\sqrt{2} & 1 & \sqrt{2} \\ 4 & 0 & 2\sqrt{2} & \sqrt{17} & \sqrt{10} \\ 2\sqrt{2} & 2\sqrt{2} & 0 & \sqrt{5} & \sqrt{10} \\ 1 & \sqrt{17} & \sqrt{5} & 0 & \sqrt{5} \\ \sqrt{2} & \sqrt{10} & \sqrt{10} & \sqrt{5} & 0 \end{pmatrix} \approx \begin{pmatrix} 0 & 4 & 2.828 & 1 & 1.414 \\ 4 & 0 & 2.828 & 4.123 & 3.162 \\ 2.828 & 2.828 & 0 & 2.236 & 3.162 \\ 1 & 4.123 & 2.236 & 0 & 2.236 \\ 1.414 & 3.162 & 3.162 & 2.236 & 0 \end{pmatrix}$$

Observe that if node 1 has a deadline of 4 units then it must be visited first in a tour to keep from violating this deadline. So we will consider the two tours above: 012340 (Tour I) and 041230 (Tour II).
Using the expected value formula from above, we find the expected cost of Tour I is (with $n = 4$):

$$\begin{aligned} E(L_\tau) = \;& \sum_{j=1}^{4} p_j d_{0,j} \prod_{k=1}^{j-1}(1-p_k) + \sum_{i=1}^{3}\sum_{j=i+1}^{4} p_i p_j d_{i,j} \prod_{k=i+1}^{j-1}(1-p_k) + \sum_{i=1}^{4} p_i d_{i,0} \prod_{k=i+1}^{4}(1-p_k) \\ = \;& p_1 d_{01} + p_2 d_{02}(1-p_1) + p_3 d_{03}(1-p_1)(1-p_2) + p_4 d_{04}(1-p_1)(1-p_2)(1-p_3) \\ & + p_1 p_2 d_{12} + p_1 p_3 d_{13}(1-p_2) + p_1 p_4 d_{14}(1-p_2)(1-p_3) \\ & + p_2 p_3 d_{23} + p_2 p_4 d_{24}(1-p_3) + p_3 p_4 d_{34} \\ & + p_4 d_{04} + p_3 d_{03}(1-p_4) + p_2 d_{02}(1-p_3)(1-p_4) + p_1 d_{01}(1-p_2)(1-p_3)(1-p_4) \\ = \;& (0.1)(4) + (1)(2.828)(0.9) + (0.1)(1)(2.828) + (1)(1)(2.236) \\ & + (1)(0.5)(2.236) + (0.5)(1.414) + (1)(1)(0.5) \\ = \;& 0.4 + 2.545 + 0.2828 + 2.236 + 1.118 + 0.707 + 0.5 \\ = \;& 7.789 \end{aligned}$$

We observe that Tour II only causes a deadline violation when both nodes 4 and 1 are active, which has a probability of $p_1 p_4 = 0.1 \cdot 0.5 = 0.05 = 5\%$.
In this case, the expected cost of $\tau = 041230$ (Tour II) is:

$$\begin{aligned} E(L_\tau) = \;& p_4 d_{04} + p_1 d_{01}(1-p_4) + p_2 d_{02}(1-p_4)(1-p_1) + p_3 d_{03}(1-p_4)(1-p_1)(1-p_2) \\ & + p_4 p_1 d_{14} + p_4 p_2 d_{24}(1-p_1) + p_4 p_3 d_{34}(1-p_1)(1-p_2) + p_1 p_2 d_{12} + p_1 p_3 d_{13}(1-p_2) \\ & + p_2 p_3 d_{23} + p_3 d_{03} + p_2 d_{02}(1-p_3) + p_1 d_{01}(1-p_3)(1-p_2) + p_4 d_{04}(1-p_3)(1-p_2)(1-p_1) \\ = \;& (0.5)(1.414) + (0.1)(4)(0.5) + (2.828)(0.5)(0.9) + (0.5)(0.1)(3.162) \\ & + (0.5)(3.162)(0.9) + (0.1)(2.828) + (2.236) + (1) \\ = \;& 0.707 + 0.2 + 1.273 + 0.158 + 1.423 + 0.283 + 2.236 + 1 \\ = \;& 7.28 \end{aligned}$$

Thus the expected value of Tour I is 7.789 while the expected value of Tour II is 7.280 but might violate the deadline constraint for node 1 at roughly a 5% chance (which may be within the expected tolerance, depending on the situation).