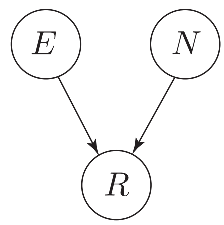
**ECE C143A/C243A, Spring 2022** **Homework #5**
Department of Electrical and Computer Engineering Prof. J.C. Kao
University of California, Los Angeles TAs T. Monsoor, W. Yu

Due Friday, 20 May 2022, uploaded to Gradescope.
Covers material up to Graphical Models.
100 points total.

1. (32 points) Let's say that you're recording from a neuron with equipment that breaks down from time to time. If you don't see a spike recorded by the equipment, you want to know whether to attribute that to the equipment being broken or to the neuron really not spiking. The following directed graph describes this scenario:




where $E$ indicates whether the neural recording equipment is working ($E = 1$) or broken ($E = 0$), $N$ indicates whether the neuron spiked ($N = 1$) or did not spike ($N = 0$), and $R$ indicates whether a spike was recorded ($R = 1$) or not recorded ($R = 0$) by the equipment.

The neuron spikes independently of whether the recording equipment is working or broken, with prior probabilities:

$$P(E = 1) = 0.9$$
$$P(N = 1) = 0.25$$

Given the values of $E$ and $N$, a spike is recorded with the following probabilites:

$$P(R = 1|E = 1, N = 1) = 1$$
$$P(R = 1|E = 1, N = 0) = 0$$
$$P(R = 1|E = 0, N = 1) = 0.1$$
$$P(R = 1|E = 0, N = 0) = 0.1$$

(a) (6 points) What is the prior probability that the neuron did not spike?

(b) (10 points) What is the probability that the neuron did not spike, given that no spike was recorded? Why does this answer make sense intuitively compared to the answer in part (a) (i.e., which is larger)?

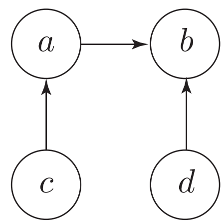
(c) (10 points) What is the probability that the neuron did not spike, given that no spike was recorded *and* that the equipment was broken? Why does this answer make sense intuitively compared to the answer in part (b) (i.e., which is larger)?

(d) (6 points) Are $E$ and $N$ independent, given $R$? (Hint: To show conditional independence, one must show that

$$P(E, N|R) = P(E|R)P(N|R) \tag{1}$$

is satisfied for all possible combinations of $E$, $N$, and $R$. However, to show conditional *dependence*, one needs to only show that (1) is not satisfied for one possible combination of $E$, $N$, and $R$.)

2. (34 points) Consider the following directed graph:




   (a) (4 points) Write the joint distribution $P(a, b, c, d)$ in terms of a product of conditional distributions, one for each node in the graph. (Hint: Use equation (8.5) in *PRML*. Note that if a node has no parents, then its conditional distribution becomes a marginal distribution.)

   (b) (6 points) Are $c$ and $d$ independent? Justify mathematically (as opposed to only an intuitive explanation).

   (c) (4 points) Are $c$ and $d$ independent, given $a$ and $b$? Justify mathematically.

   (d) (6 points) Are $a$ and $d$ independent? Justify mathematically.

   (e) (4 points) Are $a$ and $d$ independent, given $b$? Justify mathematically.

   (f) (6 points) Are $b$ and $c$ independent? Justify mathematically.

   (g) (4 points) Are $b$ and $c$ independent, given $a$? Justify mathematically.

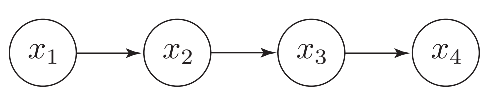
3. (14 points) In HW 4, we considered *naive Bayes* models, where the observed dimensions $x_i$ ($i = 1, \ldots, D$) are independent, conditioned on the class $z \in \{1, \ldots, K\}$.

   (a) (6 points) Draw a directed graph representing the joint probability distribution over the variables $x_1, \ldots, x_D$ and $z$. There should be one node for each of the $D + 1$ variables.

   (b) (4 points) Write the joint distribution $P(x_1, \ldots, x_D, z)$ in terms of a product of conditional distributions, one for each node in the graph.

   (c) (4 points) Are $x_1, \ldots, x_D$ independent? Why?

4. (20 points) Consider the following directed graph:




   (a) (4 points) Write the joint distribution $P(x_1, x_2, x_3, x_4)$ in terms of a product of conditional distributions, one for each node in the graph.

(b) (8 points) A possible application of this directed graph is to model a stimulus that changes over time. Let

$$x_1 \sim \mathcal{N}\left(0,\, \sigma^2\right)$$
$$x_t | x_{t-1} \sim \mathcal{N}\left(x_{t-1},\, \sigma^2\right),$$

where $t = 2, 3, 4$. This is known as a *random walk model*, since $x_t$ is obtained by adding a random increment (in this case, a Gaussian) to $x_{t-1}$. What is the $4 \times 4$ covariance matrix of the vector $[x_1\ x_2\ x_3\ x_4]^{\mathrm{T}}$ in terms of $\sigma^2$?

(c) (4 points) Find the inverse of the covariance matrix found in part (b) in terms of $\sigma^2$. This is known as the *precision matrix*. You can do this inversion in `Python`.

(d) (4 points) Relate where zeros appear in the precision matrix to the graph structure.