Homework 5: Graphical Models, MDPs

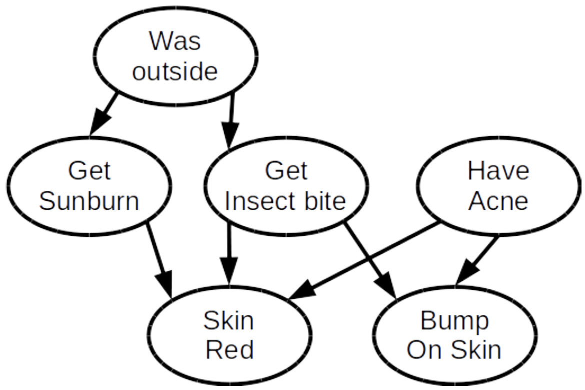
Introduction

There is a mathematical component and a programming component to this homework. Please submit your **tex, PDF, and Python files** to Canvas, and push all of your work to your GitHub repository. If a question requires you to make any plots, please include those in the writeup.

Bayesian Networks [7 pts]

Problem 1

In this problem we explore the conditional independence properties of a Bayesian Network. Consider the following Bayesian network representing a person's skin condition. Each random variable is binary (true/false).



The random variables are:

- Was Outside: Was the person outside?
- Get Sunburn: Did the person get sunburn?
- Get Insect Bite: Did the person get an insect bite?
- Have Acne: Does the person have acne?
- Skin Red: Is the skin red?
- Bump on Skin: Is there a bump?

For the following questions, $I(A, B)$ means that events A and B are independent and $I(A, B|C)$ means that events A and B are independent conditioned on C. Use the concept of d-separation to answer the questions and show your work.

1. Is $I(\text{Have Acne}, \text{Was Outside})$? If NO, give intuition for why.
2. Is $I(\text{Have Acne}, \text{Was Outside} \,|\, \text{Skin Red})$? If NO, give intuition for why.
3. Is $I(\text{Get Sunburn}, \text{Bump on Skin})$? If NO, give intuition for why.
4. Is $I(\text{Get Sunburn}, \text{Bump on Skin} \,|\, \text{Get Insect Bite})$? If NO, give intuition for why.
5. Suppose the person has taken a medicine to suppress their response to insect bites: they get red skin, but no bumps. Draw the modified network.
6. For this modified network, is $I(\text{Get Sunburn}, \text{Bump on Skin})$? If NO, give an intuition why. If YES, describe what observations (if any) would cause them to no longer be independent.

Solution Bayesian Networks

Solution: Please see figure for showing d-separation work.

- Is $I(\mathsf{Have\ Acne}, \mathsf{Was\ Outside})$? If NO, give intuition for why.

  YES, Yes, having acne is independent of going outside. Whether we go outside or not doesn't affect whether we get acne (at least according to this model). Also, having acne doesn't affect whether we go outside or not (according to this model).

- Is $I(\mathsf{Have\ Acne}, \mathsf{Was\ Outside} \mid \mathsf{Skin\ Red})$? If NO, give intuition for why.

  NO, If we know whether or not we we have acne, does also knowing whether our skin is red or not impact our estimate of of whether we were outside? If we know we have red skin, and knowing we had acne would "explain away" some of our certainty about being outside or not. Thus, having observed R, O and A are no longer independent.

- Is $I(\mathsf{Get\ Sunburn}, \mathsf{Bump\ on\ Skin})$? If NO, give intuition for why.

  NO, Not independent. The likelihood of us getting sunburned impacts our estimate of being outside, which in turn impacts our estimate of getting an insect bite and potentially having a bump on our skin.

- Is $I(\mathsf{Get\ Sunburn}, \mathsf{Bump\ on\ Skin} \mid \mathsf{Get\ Insect\ Bite})$? If NO, give intuition for why. YES, Given that we got an insect bite, do we care about whether we have a bump or not, in order to estimate if we are sunburned or not? No! (thus YES, they are independent).

- Suppose the person has taken a medicine to suppress their response to insect bites: they get red skin, but no bumps. Draw the modified network.

- For this modified network, is $I(\mathsf{Get\ Sunburn}, \mathsf{Bump\ on\ Skin})$? If NO, give an intuition why. If YES, describe what observations (if any) would cause them to no longer be independent.

$$sdf = 1 \tag{1}$$

Kalman Filters [7 pts]

Problem 2

In this problem, you will implement a one-dimensional Kalman filter. Assume the following dynamical system model:

$$\begin{aligned} z_{t+1} &= z_t + \epsilon_t \\ x_t &= z_t + \gamma_t \end{aligned}$$

where z are the hidden variables and x are the observed measurements. The random variables ϵ and γ are drawn from the following Normal distributions:

$$\begin{aligned} \epsilon_t &\sim N(\mu_\epsilon, \sigma_\epsilon) \\ \gamma_t &\sim N(\mu_\gamma, \sigma_\gamma) \end{aligned}$$

where $\mu_\epsilon = 0$, $\sigma_\epsilon = 0.05$, $\mu_\gamma = 0$ and $\sigma_\gamma = 1.0$

You are provided with the observed data x and the hidden data z in kf-data.csv, and the prior on the first hidden state is $p(z_0) = \mathrm{N}(\mu_p, \sigma_p)$ where $\mu_p = 5$ and $\sigma_p = 1$

(a) The distribution $p(z_t|x_0 ... x_t)$ will be Gaussian $N(\mu_t, \sigma_t^2)$. Derive an iterative update for the mean μ_t and variance σ_t^2 given the mean and variance at the previous time step (μ_{t-1} and σ_{t-1}^2).

(b) Implement this update and apply it to the observed data above (do not use the hidden data to find these updates). Provide a plot of μ_t over time as well as a

- Model this as a Markov decision process: define the states S, actions A, reward function $r : S \times A \mapsto \mathbb{R}$, and transition model $p(s' \mid s, a)$ for $s', s \in S$ and $a \in A$. For now, assume that the robot's actions execute perfectly: if the robot tries to move in a particular direction, it always succeeds in doing so.

- Consider a *random policy* π, where in each state the robot moves uniformly at randomly in any of its available directions (including off the board). For every position on the grid calculate the value function, $V_t^\pi : S \mapsto \mathbb{R}$, under this policy, for $t = 2, 1$ steps left to go. You can find LaTeX code for the tables in the solution template. Note that you should have 2 tables, one for each time horizon.

- Now assume that the robot plays an *optimal policy* π_t^* (for t time steps to go). Find the optimal policy in the case of a finite time horizon of $t = 1, 2$ and give the corresponding MDP value functions $V_t^* : S \mapsto \mathbb{R}$, under this optimal policy. You can indicate the optimal policy for each time horizon on the corresponding V_t^* table via arrows or words in the direction that the robot should move from that state.

- Now consider the situation where the robot does not have complete control over its movement. In particular, when it chooses a direction, there is a 80% chance that it will go in that direction, and a 10% chance it will go in the two adjacent (90° left or 90° right) directions. Explain how this changes the elements S, A, r, and $p(s' \mid s, a)$ of the MDP model. Assume the robot uses the same policy π_t^* from the previous question (now possibly non-optimal), and write this as π_t, and tie-break in favor of N, then E, then S then W. Give the corresponding MDP value functions $V_t^\pi : S \mapsto \mathbb{R}$, for this policy in this partial control world, for $t = 2, 1$ steps left to go. Is the policy still optimal?

Problem 3 (Calibration, 1pt)

Approximately how long did this homework take you to complete?

- Name:
- Email:
- Collaborators: