CS109: Probability for Computer Scientists

Lecture 12 — General Inference

Feb 2

Problem 1: Groundhog Day

Did you know today is Groundhog's Day?

Sees shadow = 6 more weeks of winter.

Doesn't see shadow = early spring.

Based on historical data:

- When the groundhog sees shadow, 70% chance of 6 more weeks of winter.
- When the groundhog doesn't see shadow, 40% chance of 6 more weeks of winter.
- Before we observe the groundhog, we think it is equally likely to be an early spring or 6 more weeks of winter.

a) Compute the probability of 6 more weeks of winter *given that the groundhog sees his shadow.*

Solution

Let W = "6 more weeks of winter" and S = "sees shadow". From the text:

$$P(W) = 0.5, \quad P(W^c) = 0.5, \quad P(W \mid S) = 0.7, \quad P(W \mid S^c) = 0.4.$$

Convert to likelihoods:

$$P(S \mid W) = 0.7, \qquad P(S \mid W^c) = 0.6$$

(since $P(S^c \mid W^c) = 0.4$ from the second bullet).
Then Bayes:

$$P(W \mid S) = \frac{P(S \mid W)P(W)}{P(S \mid W)P(W) + P(S \mid W^c)P(W^c)} = \frac{0.7 \cdot 0.5}{0.7 \cdot 0.5 + 0.6 \cdot 0.5} = \frac{0.35}{0.65} \approx 0.54.$$

So the PMF is:

$$P(W \mid S) \approx 0.54, \qquad P(W^c \mid S) \approx 0.46.$$

Problem 2: Lidar in 1D

Let T be the true distance. Your prior belief is: $T \sim N(\mu = 1,\ \sigma^2 = 3)$. Your sensor has uncertainty:

$$X \mid (T = t) \sim \mathcal{N}(\mu = t,\ \sigma^2 = 1.5).$$

You observe: $X = 4$.

a) Write Bayes' rule for the posterior density in the form

$$f(T = t \mid X = 4) \ \propto \ f(X = 4 \mid T = t) \cdot f(T = t).$$

(No need to simplify.)

b) (Optional) Compute the posterior distribution for T given $X = 4$. Specifically, give the mean and variance of the posterior distribution. Will require some algebra and completing the square.

c) **(Optional — Super challenge: 2D Tracking)**

Now the object is at an unknown location (X, Y).

Prior: you believe (X, Y) is centered around $(3, 3)$ with joint density

$$f(X = x, Y = y) = \frac{1}{8\pi} \exp\left(-\frac{(x-3)^2 + (y-3)^2}{8}\right).$$

Likelihood: you observe a noisy distance reading D from a sensor at $(0, 0)$. The sensor model is

$$D \mid (X = x, Y = y) \sim \mathcal{N}\left(\mu = \sqrt{x^2 + y^2},\ \sigma^2 = 1\right),$$

and you observe $D = 4$. Write the unnormalized posterior density using $\propto$:

$$f(X = x, Y = y \mid D = 4) \ \propto \ f(D = 4 \mid X = x, Y = y) \cdot f(X = x, Y = y).$$

Solution

a)

$$f(T = t \mid X = 4) \propto f(X = 4 \mid T = t)\, f(T = t).$$

b) Normal-normal update with prior variance $\sigma_0^2 = 3$ and sensor variance $\sigma_x^2 = 1.5$:

$$\sigma^2_{\text{post}} = \left(\frac{1}{\sigma_0^2} + \frac{1}{\sigma_x^2}\right)^{-1} = \left(\frac{1}{3} + \frac{1}{1.5}\right)^{-1} = \left(\frac{1}{3} + \frac{2}{3}\right)^{-1} = 1.$$

$$\mu_{\text{post}} = \sigma^2_{\text{post}}\left(\frac{\mu_0}{\sigma_0^2} + \frac{x}{\sigma_x^2}\right) = 1\left(\frac{1}{3} + \frac{4}{1.5}\right) = \frac{1}{3} + \frac{8}{3} = 3.$$

Therefore,

$$T \mid (X = 4) \sim \mathcal{N}(3, 1).$$

Problem 3: Size of a Joint Distribution

Suppose you have N binary random variables.

a) If $N = 9$, how many entries are in the full joint probability table?

b) For general N, how many entries are in the full joint probability table?

Solution

$$\text{Joint table size} = 2^N.$$

So for $N = 9$, $2^9 = 512$ entries.

Problem 4: Bayes Net with Probabilities

Consider the Bayes net with binary variables:

$$Flu \to Fever, \qquad (Flu, U) \to Tired,$$

where U stands for "Undergrad".

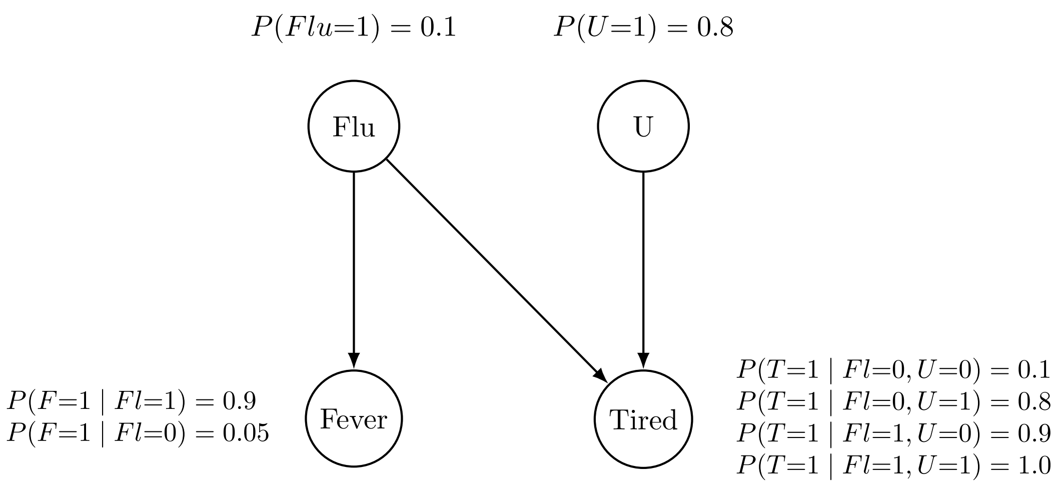
Given probabilities:

$$P(Flu = 1) = 0.1, \qquad P(U = 1) = 0.8,$$

$$P(Fever = 1 \mid Flu = 1) = 0.9, \qquad P(Fever = 1 \mid Flu = 0) = 0.05,$$

$$P(Tired = 1 \mid Flu = 0, U = 0) = 0.1$$
$$P(Tired = 1 \mid Flu = 0, U = 1) = 0.8$$
$$P(Tired = 1 \mid Flu = 1, U = 0) = 0.9$$
$$P(Tired = 1 \mid Flu = 1, U = 1) = 1.0$$

Diagram:



a) Compute $P(Fever = 0 \mid Flu = 1)$.

b) We want:

$$P(Flu = 1 \mid U = 1, Tired = 1).$$

A simulation-based estimate is:

$$P(Flu = 1 \mid U = 1, Tired = 1) \approx \frac{\#\text{ samples with } (Flu = 1, U = 1, Tired = 1)}{\#\text{ samples with } (U = 1, Tired = 1)}.$$

Explain *why* this ratio is a reasonable approximation.

Solution

Solution to part (a)

$$P(Fever = 0 \mid Flu = 1) = 1 - P(Fever = 1 \mid Flu = 1) = 1 - 0.9 = 0.1.$$

Solution

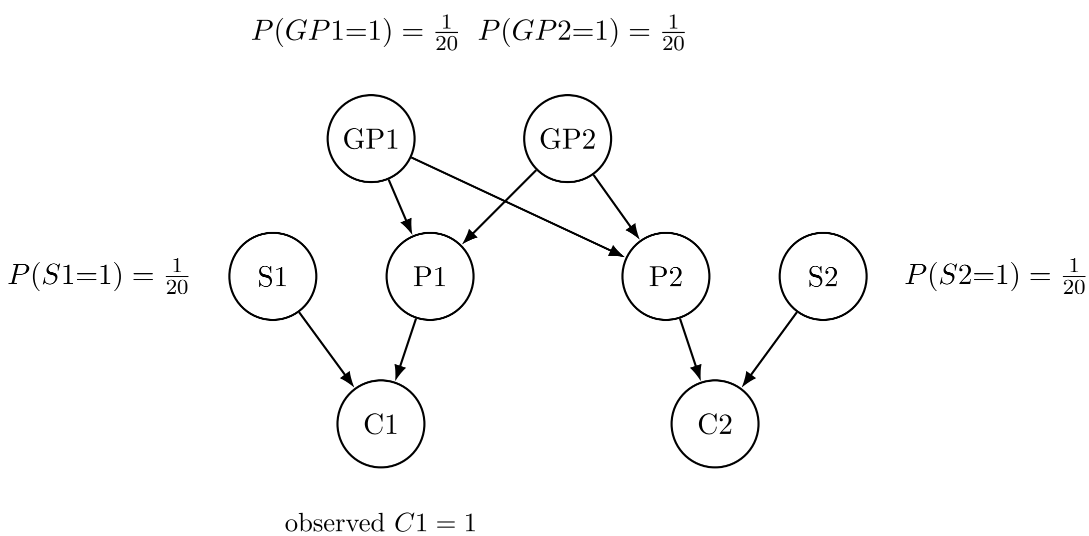
Solution to part (b)
If we generate many i.i.d. samples from the joint distribution, then (by the law of large numbers) the fraction of samples satisfying $(U = 1, Tired = 1)$ that also satisfy $Flu = 1$ approaches the true conditional probability:

$$P(A \mid B) = \lim_{n\to\infty} \frac{\#\{A \cap B\}}{\#\{B\}}.$$

Here A is $(Flu = 1)$ and B is $(U = 1, Tired = 1)$.

Problem 6: The Cousin Problem

A simplified genetic model: each person has a binary variable indicating whether they have a recessive gene (1) or not (0). We observe that **Cousin 1 has the gene**.



We want (conceptually): $P(C2 = 1 \mid C1 = 1)$.

a) Describe in words: if you do **rejection sampling**, what samples do you **throw away**?

b) Describe in words: among the samples you keep, what do you **count**?

c) What steps would you include in a function `make_sample()` that generates one full assignment for all nodes in the Bayes net?

Solution

a) Reject any sampled world where the evidence is violated; here, throw away any sample with $C1 \neq 1$ (i.e., $C1 = 0$).

b) Among accepted samples (those with $C1 = 1$), count how many have $C2 = 1$. Estimate

$$P(C2 = 1 \mid C1 = 1) \approx \frac{\#\text{accepted samples with } C2 = 1}{\#\text{accepted samples}}.$$

c) `make_sample()` does ancestral sampling: sample $GP1, GP2, S1, S2$ from priors; sample $P1, P2$ from their CPTs given $(GP1, GP2)$; sample $C1$ from its CPT given $(P1, S1)$; sample $C2$ from its CPT given $(P2, S2)$; return the full assignment.