15

Bayes' Nets: Sampling

⇒ Why Sample?

→ Learning: Get samples from a distribution you don't know.

→ Inference: Getting a sample is faster than computing the right answer.

* Sampling from given distribution

* Step1: Get sample u from uniform distribution over [0,1]

* Step 2: Convert this sample u into an outcome for the given distribution by having each target outcome associated with a sub-interval of [0,1] with sub-interval size equal to probability of the outcome.

* Sampling in Bayes' Nets

① Prior Sampling

② Rejection Sampling

③ Likelihood Weighting

④ Gibbs Sampling

★ Prior Sampling

- For $i = 1, 2, \cdots, n$
    - Sample x_i from $P(X_i \mid Parents(X_i))$
- Return $(X_1, X_2, \cdots X_n)$

⟹ This process generates samples with probability:

$$S_{PS}(x_1, \cdots x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i)) = P(x_1, \cdots x_n)$$

⟹ Let the number of samples of an event be $N_{PS}(x_1, \cdots x_n)$

⟹
$$\lim_{N \to \infty} \tilde{P}(x_1, \cdots x_n) = \lim_{N \to \infty} N_{PS}(x_1, \cdots x_n) / N$$
$$= S_{PS}(x_1, \cdots x_n)$$
$$= P(x_1, \cdots x_n)$$

⟹ I.e, the sampling procedure is consistent.

★ Rejection Sampling

Input: evidence instantiation

For $i = 1, 2, \cdots n$

- Sample x_i from $P(X_i \mid Parents(X_i))$
- If x_i not consistent with evidence
    - Reject: return - no sample is generated in this cycle

Return $(X_1, X_2 \cdots X_n)$

★ Likelihood Weighting

⇒ Problem with rejection Sampling:

↳ If evidence is unlikely, rejects lots of samples.

↳ Evidence not exploited as you sample

Idea: Fix evidence variables and sample the rest.

↳ Problem: Sample distribution not consistent.

↳ **Solution:** weight by probability of evidence given parents.

- Input: evidence instantiation
- $W = 1.0$
- for $i = 1, 2, \dots n$
  - If X_i is an evidence variable
    - ↳ X_i = Observation x_i for X_i
    - ↳ Set $w = w * P(x_i \mid Parents(X_i))$
  - else
    - ↳ Sample x_i from $P(X_i \mid Parents(X_i))$
- return $(x_1, x_2 \cdots x_n), w$

⇒ Sampling distribution if z sampled and e fixed evidence

$$S_{WS}(z,e) = \prod_{i=1}^{l} P(z_i \mid Parents(Z_i))$$

⇒ Now, samples have weights

$$w(z,e) = \prod_{i=1}^{m} P(e_i \mid Parents(E_i))$$

⇒ Together, weighted sampling distribution is consistent

$$S_{WS}(z,e) \cdot w(z,e) = \prod_{i=1}^{l} P(z_i \mid Parents(z_i)) \prod_{i=1}^{m} P(e_i \mid Parents(e_i))$$

$$= P(z,e)$$

⇒ Likelihood weighting doesn't solve all our problems:

- ↳ Evidence influences the choice of downstream variables, but not upstream ones
- ↳ We would like to consider evidence when we sample every variable (leads to Gibbs sampling)

★ Gibbs Sampling

- **Procedure**
  - → Keep track of a full instantiation $X_1, X_2, \cdots X_n$.
  - → Start with an arbitrary instantiation consistent with the evidence.
  - → Sample one variable at a time, conditioned on all the rest, but keep evidence fixed.
  - → Keep repeating this for a long time.

- **Property**
  - → In the limit of repeating this infinitely many times the resulting samples comes from the correct distribution.

- **Rationale**
  - → Both upstream & downstream variable condition on evidence.

⇒ Gibbs sampling is a special case of more general methods called Markov chain Monte Carlo (MCMC) methods.