Rejection + Likelihood Sampling

Frederik Mallmann-Trenn
6CCS3AIN

Rejection sampling

- Use Rejection sampling.
- Generate samples as before (using prior sampling)
- If the sample is such that **e** holds use it to build an estimate
- Otherwise, ignore it

Rejection sampling

function REJECTION-SAMPLING(X, **e**, bn, N) **returns** an estimate of $P(X|\mathbf{e})$

    local variables: **N**, a vector of counts over X, initially zero

    for j = 1 to N **do**
        x ← PRIOR-SAMPLE(bn)
        if x is consistent with **e then**
            N[x] ← **N**[x]+1 where x is the value of X in **x**
    return NORMALIZE(**N**[X])

Rejection sampling

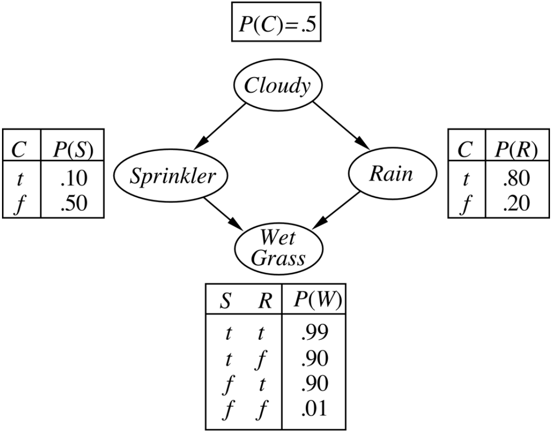
- More efficient than prior sampling, but
- For unlikely events, may have to wait a long time to get enough matching samples.
- Still inefficient.
- So, use likelihood weighting

Likelihood weighting

- Version of importance sampling.
- Fix evidence variable to $true$, so just sample relevant events.
- Have to weight them with the likelihood that they fit the evidence.
- Use the probabilities we know to weight the samples.

Likelihood weighting

- Consider we have the following network:



- Say we want to establish $\mathbf{P}(Rain|Cloudy = true, WetGrass = true)$

Likelihood weighting

- We want $\mathbf{P}(Rain|Cloudy = true, WetGrass = true)$
- We pick a variable ordering, say:

  $Cloudy$, $Sprinkler$, $Rain$, $WetGrass$.

  as before.
- Set the weight $w = 1$ and we start.
- Deal with each variable in order.

Likelihood weighting

- Remember, we want $\mathbf{P}(Rain|Cloudy = true, WetGrass = true)$
- $Cloudy$ is true, so:

$$w \leftarrow w \cdot P(Cloudy = true)$$
$$w \leftarrow 0.5$$

- $Cloudy = true$, $Sprinkler =?$, $Rain =?$, $WetGrass =?$.
- $w = 0.5$

Likelihood weighting

- $Sprinkler$ is not an evidence variable, so we don't know whether it is true or false.
- Sample a value just as we did for prior sampling:

$$\mathbf{P}(Sprinkler|Cloudy = true) = \begin{pmatrix} 0.1, \\ 0.9 \end{pmatrix}$$

- Let's assume this returns $false$.
- w remains the same.
- $Cloudy = true$, $Sprinkler = false$, $Rain =?$, $WetGrass =?$.
- $w = 0.5$

Likelihood weighting

- $Rain$ is not an evidence variable, so we don't know whether it is true or false.
- Sample a value just as we did for prior sampling:

$$\mathbf{P}(Rain|Cloudy = true) = \begin{pmatrix} 0.8, \\ 0.2 \end{pmatrix}$$

- Let's assume this returns $true$.
- w remains the same.
- $Cloudy = true$, $Sprinkler = false$, $Rain = true$, $WetGrass =?$.
- $w = 0.5$

Likelihood weighting

- $WetGrass$ is an evidence variable with value $true$, so we set:

$$w \leftarrow w \cdot P(WetGrass = true | Sprinkler = false, Rain = true)$$
$$w \leftarrow 0.45$$

- $Cloudy = true$, $Sprinkler = false$, $Rain = true, WetGrass = true.$
- $w = 0.45$

Likelihood weighting

- So we end with the event $[true, false, true, true]$ and weight 0.45.
- To find a probability we tally up all the relevant events, weighted with their weights.
- The one we just calculated would tally under

$$Rain = true$$

- As before, more samples means more accuracy.

Likelihood weighting

function LIKELIHOOD-WEIGHTING(X, $\mathbf{e}$, bn, N) **returns** an estimate of $P(X|\mathbf{e})$
  local variables: $\mathbf{W}$, a vector of weighted counts over X, initially zero

  for j = 1 to N **do**
    $\mathbf{x}, w \leftarrow$ WEIGHTED-SAMPLE(bn)
    $\mathbf{W}[x] \leftarrow \mathbf{W}[x] + w$ where x is the value of X in $\mathbf{x}$
  return NORMALIZE($\mathbf{W}[X]$)

Likelihood weighting

function WEIGHTED-SAMPLE(*bn*, **e**) **returns** an event and a weight

x ← an event with n elements; w ← 1
for i = 1 **to** n **do**
  if X_i has a value x_i in **e**
    then $w \leftarrow w \times \ P(X_i = \ x_i \mid parents(X_i))$
    else x_i ← a random sample from $\mathbf{P}(X_i \mid parents(X_i))$
return x, w