Stochastic Simulation

- Idea: probabilities $\leftrightarrow$ samples
- Get probabilities from samples:

X	$count$
x_1	n_1
$\vdots$	$\vdots$
x_k	n_k
$total$	m

$\leftrightarrow$

X	$probability$
x_1	n_1/m
$\vdots$	$\vdots$
x_k	n_k/m

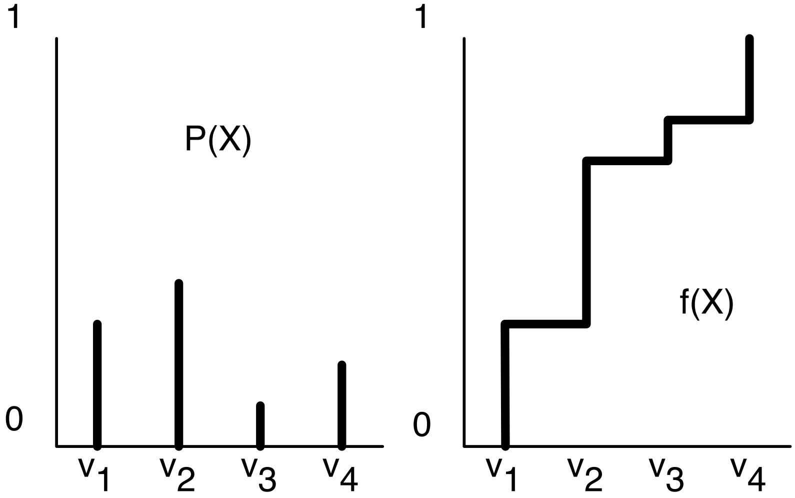
- If we could sample from a variable's (posterior) probability, we could estimate its (posterior) probability.

Generating samples from a distribution

For a variable X with a discrete domain or a (one-dimensional) real domain:

- Totally order the values of the domain of X.
- Generate the cumulative probability distribution: $f(x) = P(X \leq x)$.
- Select a value y uniformly in the range $[0, 1]$.
- Select the x such that $f(x) = y$.

Cumulative Distribution



Forward sampling in a belief network

- Sample the variables one at a time; sample parents of X before sampling X.
- Given values for the parents of X, sample from the probability of X given its parents.

Rejection Sampling

- To estimate a posterior probability given evidence $Y_1 = v_1 \wedge \ldots \wedge Y_j = v_j$:
- Reject any sample that assigns Y_i to a value other than v_i.
- The non-rejected samples are distributed according to the posterior probability:

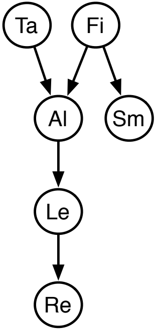
$$P(\alpha|evidence) \approx \frac{\sum_{sample \models \alpha} 1}{\sum_{sample} 1}$$

where we consider only samples consistent with evidence.

Rejection Sampling Example: $P(ta|sm, re)$

Observe $Sm = true, Re = true$

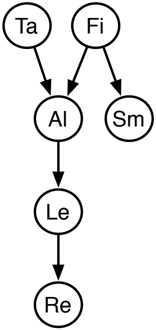
	Ta	Fi	Al	Sm	Le	Re
s_1	false	true	false	true	false	false



Rejection Sampling Example: $P(ta|sm, re)$

Observe $Sm = true, Re = true$

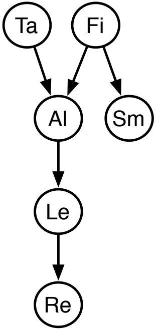
	Ta	Fi	Al	Sm	Le	Re	
s_1	false	true	false	true	false	false	✘
s_2	false	true	true	true	true	true	



Rejection Sampling Example: $P(ta|sm, re)$

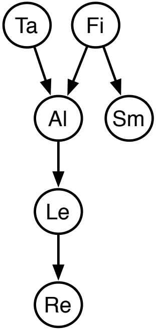
Observe $Sm = true, Re = true$

	Ta	Fi	Al	Sm	Le	Re	
s_1	false	true	false	true	false	false	✘
s_2	false	true	true	true	true	true	✔
s_3	true	false	true	false			



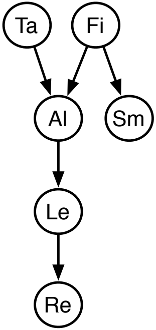
Rejection Sampling Example: $P(ta|sm, re)$

Observe $Sm = true, Re = true$



	Ta	Fi	Al	Sm	Le	Re	
s_1	false	true	false	true	false	false	✘
s_2	false	true	true	true	true	true	✔
s_3	true	false	true	false	—	—	✘
s_4	true	true	true	true	true	true	

Rejection Sampling Example: $P(ta|sm, re)$

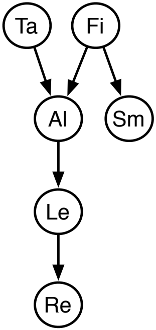


Observe $Sm = true, Re = true$

	Ta	Fi	Al	Sm	Le	Re	
s_1	false	true	false	true	false	false	✘
s_2	false	true	true	true	true	true	✔
s_3	true	false	true	false	—	—	✘
s_4	true	true	true	true	true	true	✔
...							
s_{1000}	false	false	false	false			

Rejection Sampling Example: $P(ta|sm, re)$

Observe $Sm = true$, $Re = true$



	Ta	Fi	Al	Sm	Le	Re	
s_1	false	true	false	true	false	false	✘
s_2	false	true	true	true	true	true	✔
s_3	true	false	true	false	—	—	✘
s_4	true	true	true	true	true	true	✔
...							
s_{1000}	false	false	false	false	—	—	✘

$P(sm) = 0.02$
$P(re|sm) = 0.32$
How many samples are rejected?
How many samples are used?

Importance Sampling

- Samples have weights: a real number associated with each sample that takes the evidence into account.
- Probability of a proposition is weighted average of samples:

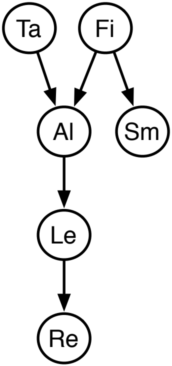
$$P(\alpha|evidence) \approx \frac{\sum_{sample \models \alpha} weight(sample)}{\sum_{sample} weight(sample)}$$

- Mix exact inference with sampling: don't sample all of the variables, but weight each sample according to $P(evidence|sample)$.

Importance Sampling (Likelihood Weighting)

procedure *likelihood_weighting*(Bn, e, Q, n):
    $ans[1:k] \leftarrow 0$ where k is size of $dom(Q)$
    repeat n times:
        $weight \leftarrow 1$
        for each variable X_i in order:
            if $X_i = o_i$ is observed
                $weight \leftarrow weight \times P(X_i = o_i | parents(X_i))$
            else assign X_i a random sample of $P(X_i | parents(X_i))$
        if Q has value v:
            $ans[v] \leftarrow ans[v] + weight$
    return $ans / \sum_v ans[v]$

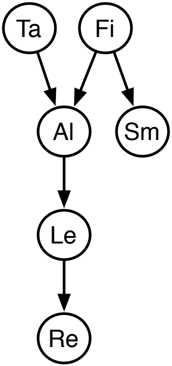
Importance Sampling Example: $P(ta|sm, re)$



	Ta	Fi	Al	Le	Weight
s_1	true	false	true	false	
s_2	false	true	false	false	
s_3	false	true	true	true	
s_4	true	true	true	true	
...					
s_{1000}	false	false	true	true	

$$P(sm|fi) = 0.9$$
$$P(sm|\neg fi) = 0.01$$
$$P(re|le) = 0.75$$
$$P(re|\neg le) = 0.01$$

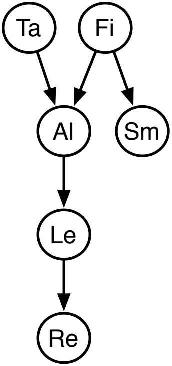
Importance Sampling Example: $P(ta|sm, re)$



	Ta	Fi	Al	Le	Weight
s_1	true	false	true	false	0.01×0.01
s_2	false	true	false	false	0.9×0.01
s_3	false	true	true	true	0.9×0.75
s_4	true	true	true	true	0.9×0.75
...					
s_{1000}	false	false	true	true	0.01×0.75

$P(sm|fi) = 0.9$
$P(sm|\neg fi) = 0.01$
$P(re|le) = 0.75$
$P(re|\neg le) = 0.01$

Importance Sampling Example: $P(le|sm, ta, \neg re)$



$$P(ta) = 0.02$$
$$P(fi) = 0.01$$
$$P(al|fi \wedge ta) = 0.5$$
$$P(al|fi \wedge \neg ta) = 0.99$$
$$P(al|\neg fi \wedge ta) = 0.85$$
$$P(al|\neg fi \wedge \neg ta) = 0.0001$$
$$P(sm|fi) = 0.9$$
$$P(sm|\neg fi) = 0.01$$
$$P(le|al) = 0.88$$
$$P(le|\neg al) = 0.001$$
$$P(re|le) = 0.75$$
$$P(re|\neg le) = 0.01$$

Particle Filtering

- Suppose the evidence is $e_1 \wedge e_2$
  $P(e_1 \wedge e_2|sample) = P(e_1|sample)P(e_2|e_1 \wedge sample)$
- After computing $P(e_1|sample)$, we may know the sample will have an extremely small probability.
- Idea: we use lots of samples: "particles". A particle is a sample on some of the variables.
- Based on $P(e_1|sample)$, we resample the set of particles. We select from the particles according to their weight.
- Some particles may be duplicated, some may be removed.

Particle Filtering for HMMs

- Start with a number of random chosen particles (say 1000)
- Each particle represents a state, selected in proportion to the initial probability of the state.
- Repeat:
  - Absorb evidence: weight each particle by the probability of the evidence given the state represented by the particle.
  - Resample: select each particle at random, in proportion to the weight of the sample.
    Some particles may be duplicated, some may be removed.
  - Transition: sample the next state for each particle according to the transition probabilities.

To answer a query about the current state, use the set of particles as data.