Sampling Based Inference

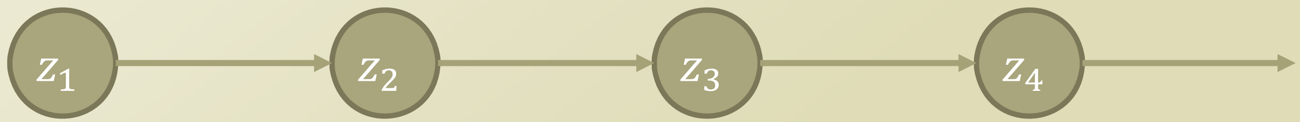
Il-Chul Moon
Dept. of Industrial and Systems Engineering
KAIST
icmoon@kaist.ac.kr

Weekly Objectives

- Learn basic sampling methods
  - Understand the concept of Markov chain Monte Carlo
  - Able to apply MCMC to the parameter inference of Bayesian networks
  - Know the mechanism of rejection sampling
  - Know the mechanism of importance sampling
- Learn sampling based inference
  - Understand the concept of Metropolis-Hastings Algorithm
  - Know the mechanism of Gibbs sampling
- Know a case study of sampling based inference
  - Understand the latent Dirichlet allocation model
  - Know the collapsed Gibbs sampling
  - Know how to derive Gibbs sampling formula for LDA

Markov Chain for Sampling

- Problem of the previous samplings?
  - No use of the past records → every sampling is independent
- Assigning Z values is a key in the inference
  - Let's assign the values by sampling result
    - Calculate P(E|MC=T,A=F) → Toss a biased coin to assign a value to E
- Sequence of random variables such a process moves through, with the Markov property defining serial dependence only between adjacent periods (as in a "chain")
- A Markov chain is a stochastic process with the Markov property
  - Example) First-order Markov chain



    - $p\left(\boldsymbol{z}^{(m+1)}\,|\boldsymbol{z}^{(1)},\ldots,\boldsymbol{z}^{(m)}\right) = p\left(\boldsymbol{z}^{(m+1)}\mid\boldsymbol{z}^{(m)}\right), m \in \{1,\ldots,M-1\}$
- Describing systems that follow a chain of linked events, where what happens next depends only on the current state of the system

Markov chain theory vs. Markov Chain Monte Carlo

- Traditional Markov Chain analysis :
  - A transition rule, $p\left(z^{(t+1)} \mid z^{(t)}\right)$, is given,
  - Interested in finding the stationary distribution $\pi(z)$

$$z_1 \rightarrow z_2 \rightarrow z_3 \rightarrow z_4 \rightarrow$$

- Markov chain Monte Carlo(MCMC) :
  - A target stationary distribution $\pi(z)$ is known,
  - Interested in prescribing an efficient transition rule to reach the stationary distribution
  - Algorithms for sampling from probability distributions based on constructing a Markov chain that has the desired distribution $\pi(z)$
  - Starting from an arbitrary state, the Markov chain proceeds
    - $\underbrace{z^{(1)} \rightarrow z^{(2)} \rightarrow \cdots \rightarrow z^{(m)}}_{Burn-in\ period} \rightarrow \underbrace{z^{(m+1)} \rightarrow z^{(m+2)} \rightarrow \cdots \rightarrow z^{(m+n)}}_{Treat\ them\ as\ samples\ from\ \pi(x)}$

Markov Chain of Z