AI – Machine Learning

Artificial Intelligence
Research Group

Posterior Distribution

- Markov Chain Monte Carlo (MCMC)
- Variational Inference

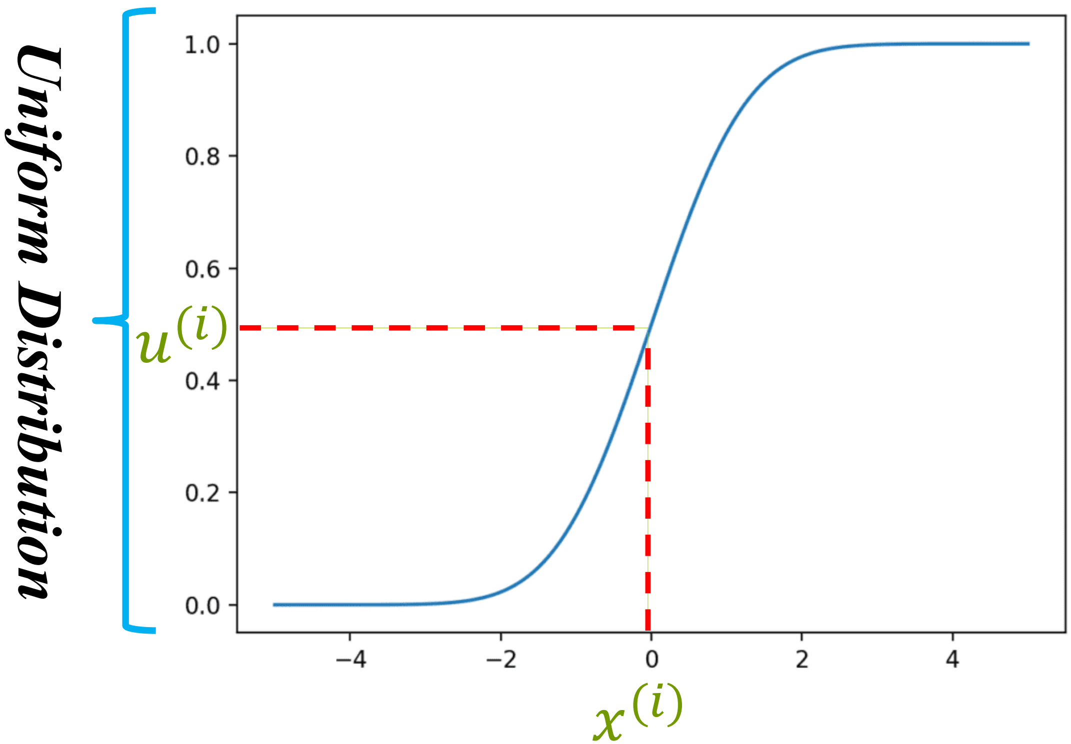
$$p(\theta|x) \propto \underbrace{p(x|\theta)p(\theta)}$$

$$\mathcal{N}(\mu|\hat{\mu}_0, \hat{\sigma}_0^2) \propto \mathcal{N}(x|\mu, \sigma^2)\mathcal{N}(\mu|\mu_0, \sigma_0^2)$$

Conjugate distribution

$$\hat{\theta} = \frac{1}{N}\sum_{i=1}^{N}\theta^{(i)}$$

Cumulative Distribution Function

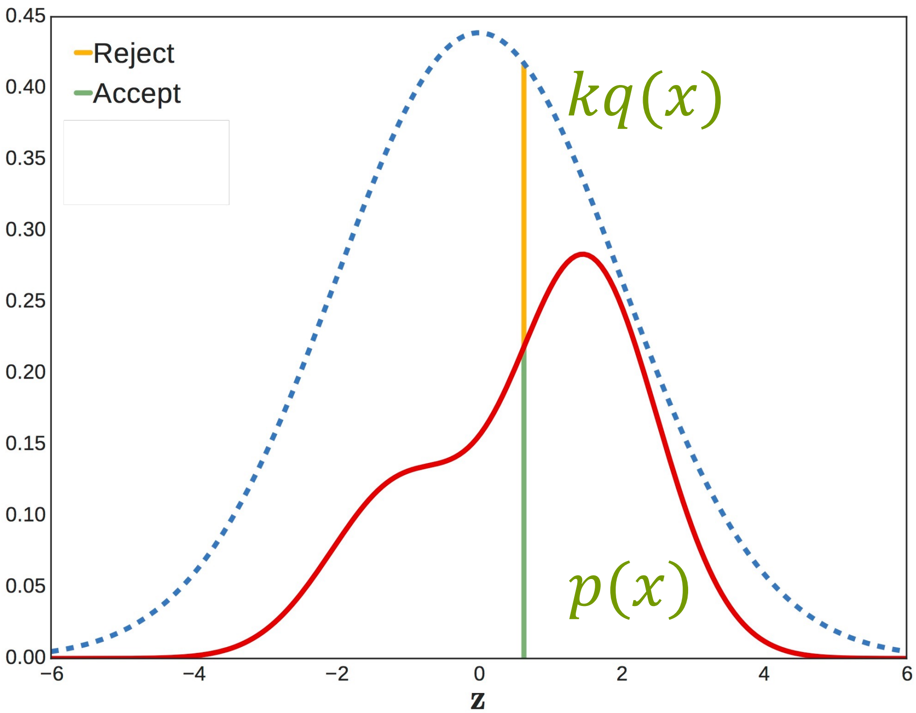


$$\mathrm{cdf}(x) = \int_{-\infty}^{t} p(t)dt$$

$$u \sim \mathcal{U}(0,1)$$

$$x \sim \mathrm{cdf}^{-1}(u)$$

Rejection Sampling



While $i \neq N$

$x^{(i)} \sim q(x)$ and $u \sim \mathcal{U}(0,1)$

if $u < \dfrac{p(x^{(i)})}{kq(x^{(i)})}$ **then**

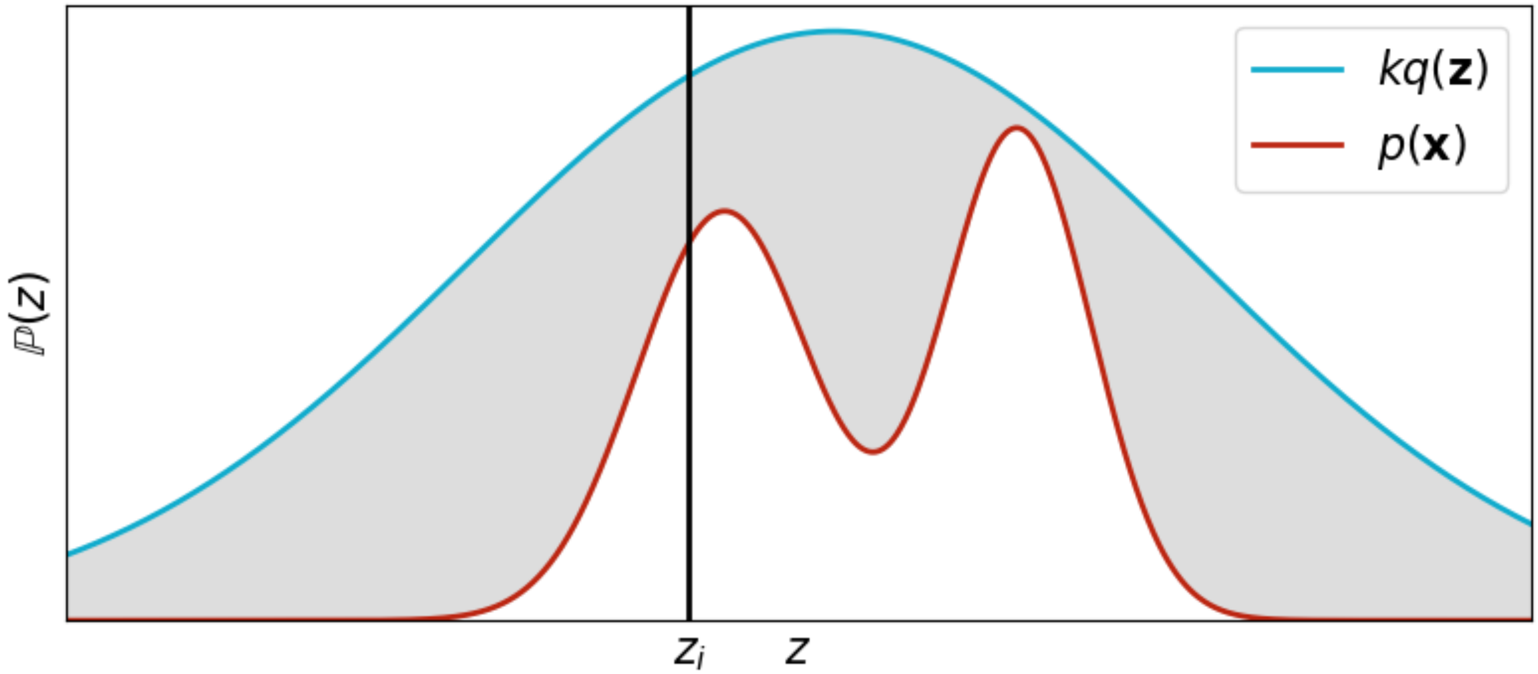
accept $x^{(i)}$

$i = i + 1$

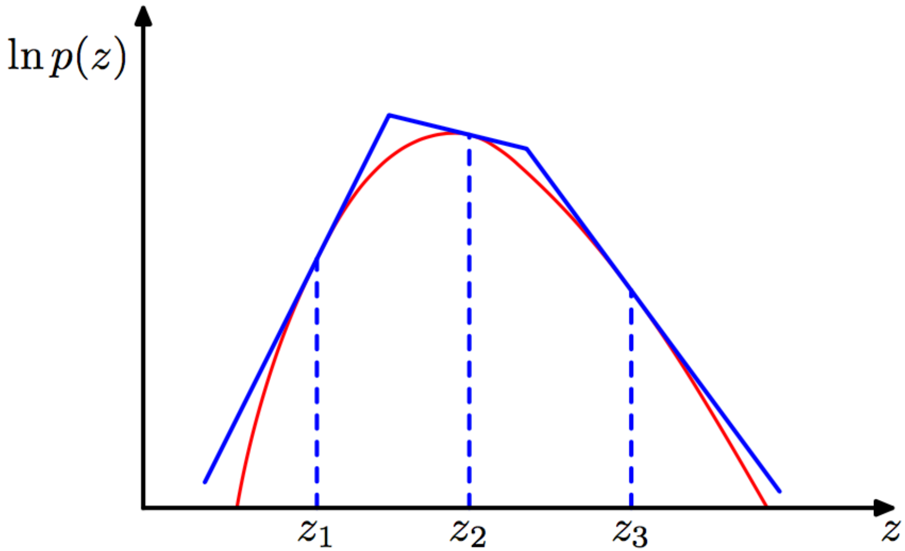
else

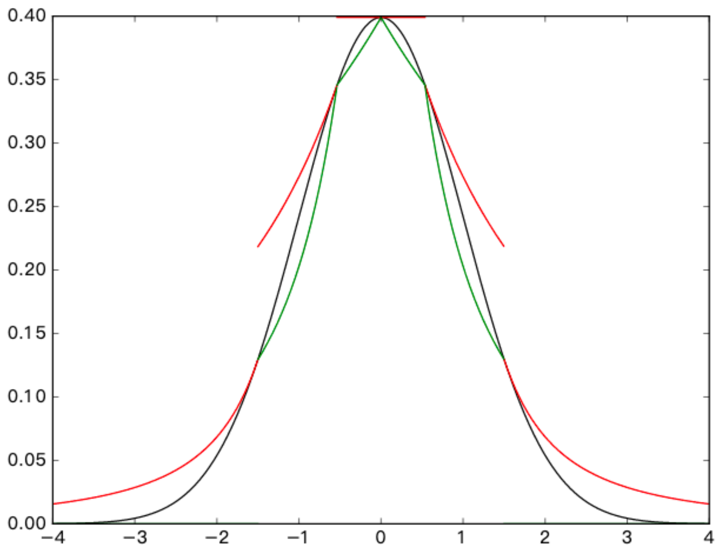
reject $x^{(i)}$

Efficiency of sampling



Piece-wise Exponential Distribution





The distribution $p(x)$ is log-concave function

Importance Sampling

$$\mathrm{E}_{p(x)}[f(x)] = \int_x f(x)p(x)dx \qquad x \sim p(x)$$

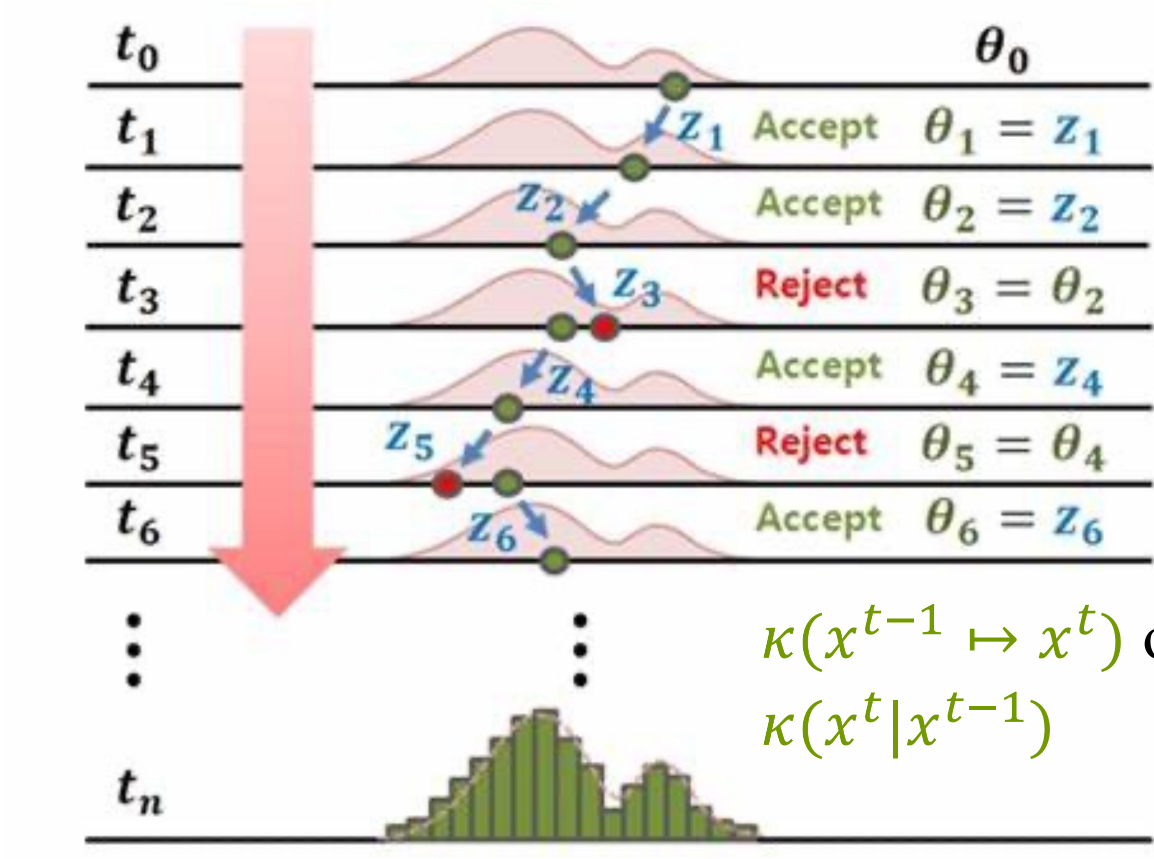
$$\hat{\mathrm{E}}_{p(x)}[f(x)] = \frac{1}{N}\sum_{i=1}^{N} f(x^{(i)})$$

$$\mathrm{E}[f(x)] = \int_x \frac{f(x)p(x)}{q(x)} q(x)dx \quad x \sim q(x)$$

$$\hat{\mathrm{E}}_{q(x)}[f(x)] = \frac{1}{N}\sum_{i=1}^{N} f(x^{(i)}) \underbrace{\frac{p(x^{(i)})}{q(x^{(i)})}}_{\textit{Weight}}$$

Markov Chain Monte Carlo

$\pi(x)$



$\kappa(x^{t-1} \mapsto x^t)$ or

$\kappa(x^t | x^{t-1})$

Chapman-Kolmogorov Equation

$$\pi_t(x^*) = \int_x \pi_{t-1}(x)\kappa(x^*|x)\, dx$$

$$p(y) = \int_x \underbrace{p(x)p(y|x)}_{p(x,y)}\, dx$$

$$\pi(x^*) = \int_x \pi(x)\kappa(x^*|x)\, dx$$ since $\pi_t(x^*) = \pi_{t-1}(x)$

not necessarily $\downarrow$ $\uparrow$ *imply*

$$\pi(x)\kappa(x^*|x) = \pi(x^*)\kappa(x|x^*)$$ ***The detailed balance***

$$\underbrace{\int_x \pi(x)\kappa(x^*|x)dx}_{} = \underbrace{\int_x \pi(x^*)\kappa(x|x^*)dx}_{= \pi(x^*)} = \pi(x^*) \underbrace{\int_x \kappa(x|x^*)dx}_{= 1}$$

Metropolis-Hastings Sampling

Initialize $x^{(0)}$

For $i = 0$ **to** $N - 1$

$x^* \sim q(x^*|x^{(i)})$ **and** $u \sim \mathcal{U}(0,1)$

if $u < \alpha(x^*) = \min\left(1, \frac{\pi(x^*)q(x|x^*)}{\pi(x)q(x^*|x)}\right)$ **then**

$x^{(i+1)} = x^*$

else

$x^{(i+1)} = x^{(i)}$

Metropolis-Hastings Sampling

$\pi(x)\kappa(x^*|x) = \pi(x^*)\kappa(x|x^*)$ ***The detailed balance***

$$\pi(x)\, q(x^*|x) \cdot \min\left(1, \frac{\pi(x^*)q(x|x^*)}{\pi(x)q(x^*|x)}\right)$$

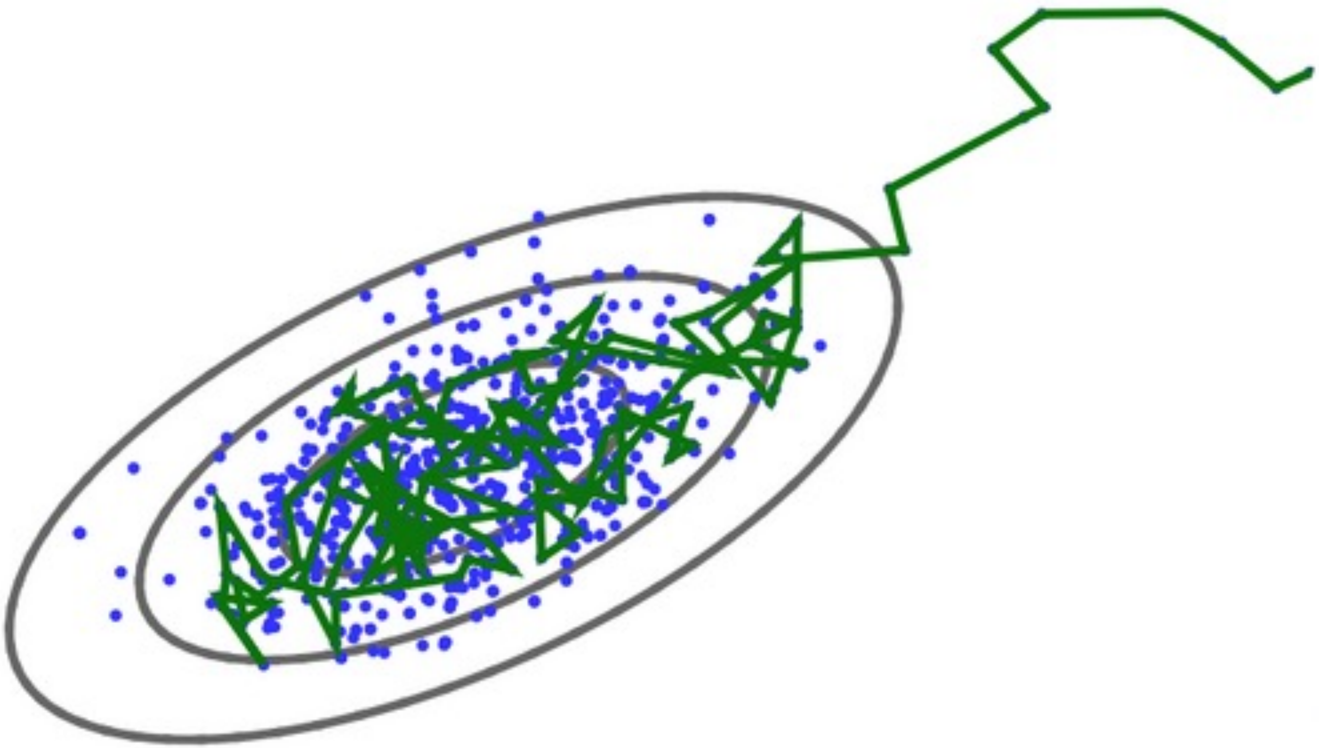
$$= \min(\pi(x)q(x^*|x), \pi(x^*)q(x|x^*))$$

$$= \pi(x^*)q(x|x^*)\min\left(\frac{\pi(x)q(x^*|x)}{\pi(x^*)q(x|x^*)}, 1\right)$$

$$= \pi(x^*)\kappa(x|x^*)$$

A good choice of $q(x^*|x)$: $\dfrac{\partial \log p(x)}{\partial x} = \dfrac{p'(x)}{p(x)}$

An Example of MH Sampling



Gibbs Sampling

$f(x, y, z)$

$X^{(1)} = (x^{(1)}, y^{(1)}, z^{(1)})$

$x^{(2)} = p(x|y^{(1)}, z^{(1)})$

$y^{(2)} = p(y|x^{(2)}, z^{(1)})$

$z^{(2)} = p(z|x^{(2)}, y^{(2)})$

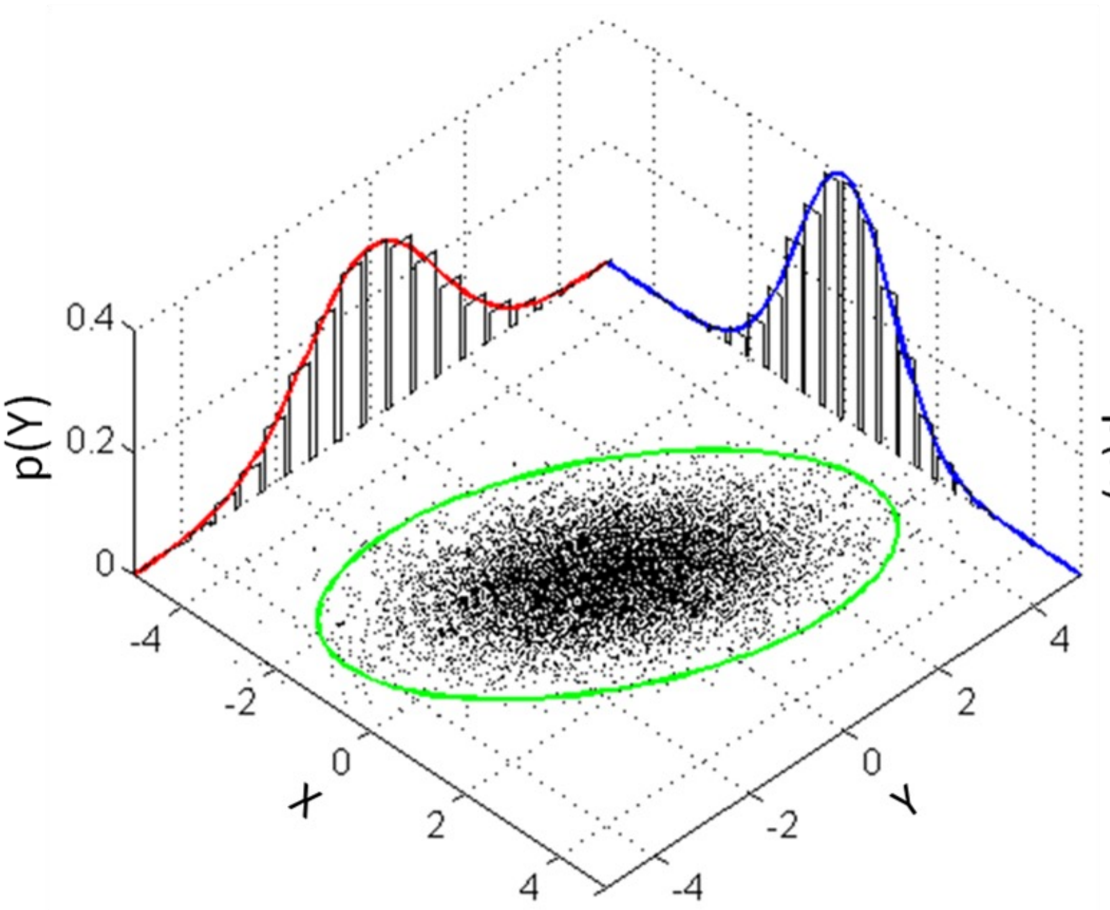
$f(x, y, z, u)$

$u \sim p(u|x, y, z)$

$p(x, y, z|u)$

$= p(x|u)p(y|u)p(z|u)$

An Example of Gibbs Sampling



Any question?

AI Research Group
Fudan University