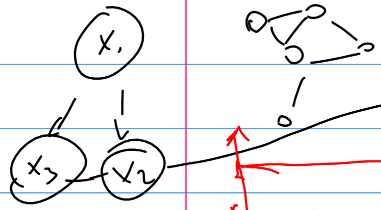
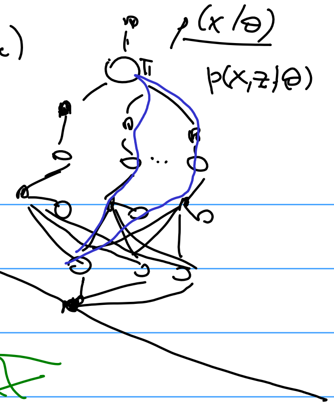
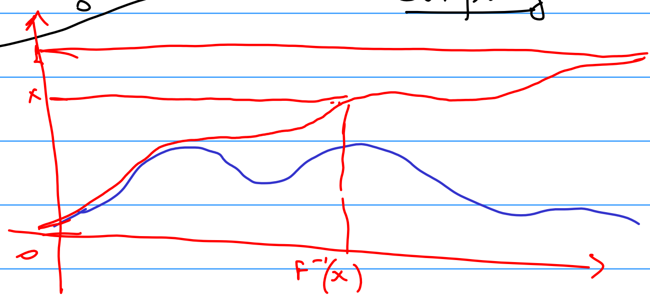


$f(X) = \prod_{x_i \subseteq X} f_i(x_i)$ $\qquad \sum_{x_1 \ldots x_n} p(x_1 \ldots x_n) = p(x_k)$

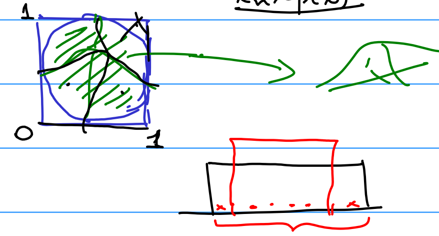
$p(x|\theta)$ $\qquad p(x,z|\theta)$



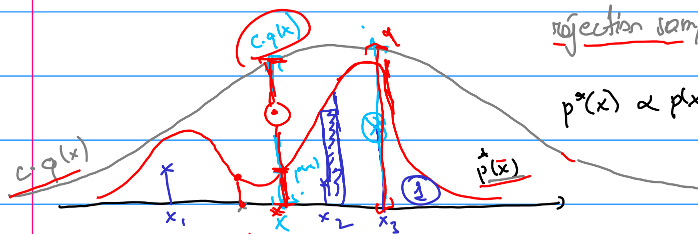
# Sampling



$\int f(x)p(x)dx \approx \frac{1}{N}\sum f(x_n)$, $\quad x_n \sim p(x)$



rejection sampling

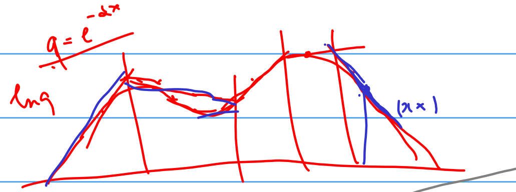


$p^*(x) \propto p(x)$

$c\,q(x) \ge p^*(x)$

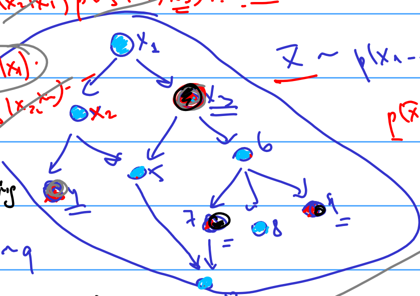
$q = e^{-\lambda x}$

$\ln q$



$\bar{x} \sim q(\bar{x}) = p(x_1)p(x_2|x_1)p(x_5|x_2,x_3)\ldots$

$p(\bar{x}) = p(x_1)p(x_2|x_1)p(x_3|x_1)\ldots p(x_4|x_2)p(x_5|x_2,x_3)\ldots$



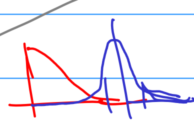
$\bar{X} \sim p(x_1 \ldots x_{10}) = p(x_1)p(x_2|x_1)\ldots p(x_{10}|x_5,x_7)$

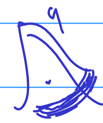
$p(\bar{x}|x_4,x_7,x_9) \propto p(\bar{x}, x_4, x_7, x_9)$

importance sampling

$E_p f \approx \frac{1}{N}\sum_{x_n \sim p(x)} f(x_n)$

$\int f(x)p(x)dx = \int f(x)\frac{p(x)}{q(x)}q(x)\,dx \approx \frac{1}{N}\sum_{x_n \sim q(x)} f(x_n)\frac{p(x_n)}{q(x_n)}$



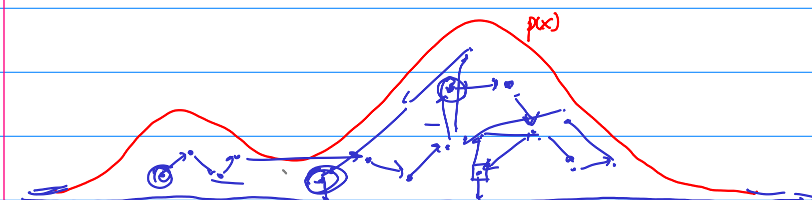


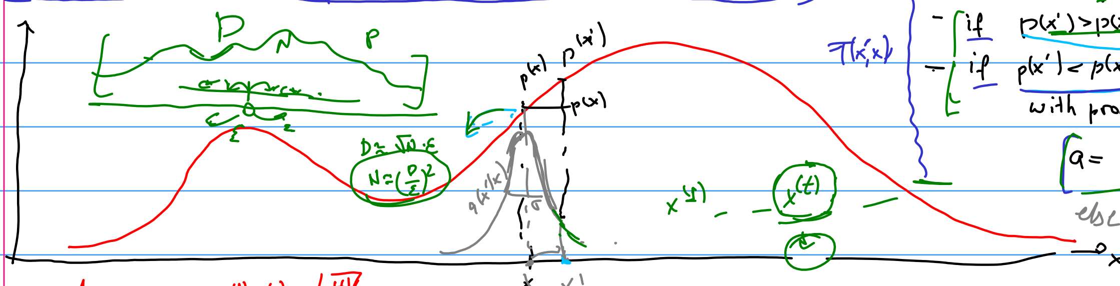
likelihood weighted sampling

$\bar{X}, x_4, x_7, x_9 \sim q$

$f(\bar{x}) \cdot \prod_{x \in \text{evidence}} p(x|par(x))$

# Markov Chain Monte Carlo (MCMC)





Metropolis-Hastings:

repeat:
- x → sampling
- sample $x' \sim q(x'|x)$
- $T(x',x)$: if $p(x') > p(x)$ accept $x'$
- if $p(x') < p(x)$ accept $x'$ with prob.

$a = \frac{p(x')}{p(x)} \cdot \frac{q(x|x')}{q(x'|x)}$

else $x' := x$

Gen. perm. $p(x)$

$A\pi = \pi$ $\qquad x^{(1)}, x^{(2)}, \ldots, x^{(t)} \ldots$ $\quad t \to \infty$ $\quad x^{(t)} \sim \pi$

Ур. баланса: $\int T(x,x')p(x')dx' = \int T(x',x)p(x)dx' \quad \forall x, x'$

$p(x) \stackrel{?}{=} \int T(x,x')p(x')dx' = p(x)\underbrace{\int T(x',x)dx'}_{=1}$

$T(x,x')p(x') = T(x',x)p(x)$

$p(x') > p(x)$: $\qquad q(x'|x)p(x)$

$\cancel{q(x|x')} \cdot \frac{p(x)}{\cancel{p(x')}} \frac{q(x'|x)}{\cancel{q(x|x')}}$

$a$

# Gibbs sampling

$\bar{x} = (x_1, x_2, \dots x_n)$  $\bar{x} \sim p(\bar{x})$,

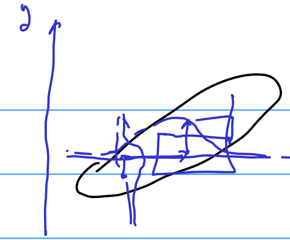
$p(x_i \mid x_1, \dots \cancel{x_i} \dots x_n) = \boxed{p(x_i \mid \bar{x}_{-i})}$

$\bar{x}^{(0)}$

$x_1 \sim p(x_1 \mid \bar{x}^{(0)}_{-1})$

$x_2 \sim p(x_2 \mid x_1, \bar{x}^{(0)}_{-1,2})$

$x_3 \sim p(x_3 \mid x_1, x_2, \dots^{(0)})$



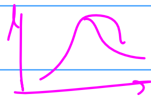
$\mathcal{N}\left(\binom{z_1}{z_2} \mid \mu, \Lambda\right) \approx q_1(z_1) \cdot q_2(z_2)$

$KL(q \| p) \to \min$  $q_1(z_1) = \mathcal{N}(z_1 \mid m_1, \Lambda_{11}^{-1})$

$q_2 = \dots \dots m_2, \Lambda_{22}^{-1}$

$\Lambda = \begin{pmatrix} \Lambda_{11} & \\ & \Lambda_{22} \end{pmatrix}$

$KL(p \| q) \to \min$



$\bar{x} = (x_1, x_2, \dots, x_i, \dots, x_n)$

$\downarrow$ $\boxed{x_i' \sim p(x_i' \mid \bar{x}_{-i})}$

$\bar{x}' = (x_1, x_2, \dots x_i', \dots, x_n)$

$q(\bar{x}' \mid \bar{x})$

$p(\bar{x}') > p(\bar{x}) \to$ accept $x'$

$p(\bar{x}') < p(\bar{x}) \to$ accept w/ pr.

$1 = A = \dfrac{p(\bar{x}')}{p(\bar{x})} \cdot \dfrac{q(\bar{x} \mid \bar{x}')}{q(\bar{x}' \mid \bar{x})} =$

$p(\bar{x}') = p(x_i' \mid \bar{x}_{-i}) p(\bar{x}_{-i})$

$" p(x_i \mid \bar{x}_{-i}) p(\bar{x}_{-i})$

$= \dfrac{\cancel{p(\bar{x}_{-i})} \, p(x_i' \mid \bar{x}_{-i})}{\cancel{p(\bar{x}_{-i})} \, p(x_i \mid \bar{x}_{-i})} \cdot \dfrac{p(x \mid \cancel{\bar{x}_{-i}})}{p(x' \mid \cancel{\bar{x}_{-i}})} = 1$

---

$p(\bar{x}) \approx q(\bar{x})$

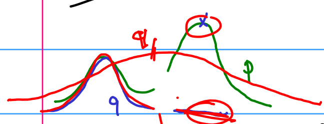
$J(p, q)$

$J(f) = \int f(x) p(x) dx \to \min$ ($+\delta f$)

EM: $p(x \mid \theta) \xrightarrow[\theta]{} \max$ | $p(x, z \mid \theta) \xrightarrow[\theta]{} \max$

$m_1 = \mu_1 - \Lambda_{11}^{-1} \Lambda_{12} (\cancel{m_2} - \mu_2)$

$m_2 = \mu_2 - \Lambda_{22}^{-1} \Lambda_{21} (\cancel{m_1} - \mu_1)$



$p(x, z \mid \theta) = p(z \mid x, \theta) p(x \mid \theta)$

$\ln p(x, z \mid \theta) = \ln p(z \mid x, \theta) + \ln p(x \mid \theta)$

$E_{q(z)}[\ln p(x \mid \theta)] = E_{q(z)}[\ln p(x, z \mid \theta) - \ln p(z \mid x, \theta)]$  $\pm \int \ln q \, q \, dz$

Kullback-Leibler

$\boxed{KL(q \| p)} = -\int \ln \frac{p}{q} q \, dx$ ($\geq 0$)

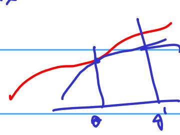
$= E_q \ln \frac{q}{p}$

$\ln p(x \mid \theta) = \int \ln p(x, z \mid \theta) q(z) dz - \int \ln p(z \mid x, \theta) q(z) dz =$

$\max_\theta \ln p(x \mid \theta) = \underbrace{\int \ln \frac{p(x, z \mid \theta)}{q(z)} q(z) dz}_{\mathcal{L}\ \max} \underbrace{- \int \ln \frac{p(z \mid x, \theta)}{q(z)} q(z) dz}_{= KL(q \| p)}$

$KL(q \| p) \geq 0$

$KL(q \| p) = 0 \Leftrightarrow q = p$



$= \boxed{\mathcal{L}(q \mid \theta)} + \underbrace{KL(q \| p(z \mid x, \theta))}_{\geq 0}$  $\uparrow q$  $\downarrow q$

$q(z) \overset{E}{=} p(z \mid x, \theta)$

$KL(p \| q) = -\int \ln \frac{q}{p} p \, dx$

$\ln p(x \mid \theta) \geq \mathcal{L}(q \mid \theta) \xrightarrow[\theta]{M} \max$

$z = (z_1, z_2, \dots z_k)$

$p(x \mid \theta) \xrightarrow[\theta]{} \max$

---

$p(x, z)$  $\boxed{p(z \mid x) \approx q(z)}$

$\mathcal{L}(q) = \int \ln \frac{p(x, z)}{q(z)} q(z) dz \to \max$

$\ln q_i^* = E_{z_{-i}} \ln p(x, z)$

$p(x, z) = p(x) p(z \mid x)$

$\ln p(x, z) - \ln p(z \mid x) = \ln p(x)$

$\prod q_i \approx p(z \mid x)$

$q(z) = \prod q_i(z_i)$

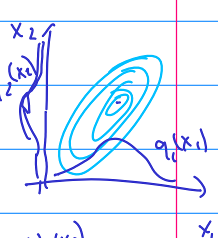
$z_i \cap z_j = \varnothing$

$\underbrace{\int \ln \frac{p(x, z)}{q(z)} q(z) dz - \int \ln \frac{p(z \mid x)}{q(z)} q(z) dz}_{} = \ln p(x)$

$\int \ln \frac{p(x, z)}{\prod q_i} \cdot \prod q_i \, dz_1 \dots dz_n = \quad -\sum_i \int (\ln q_i) \prod q_i \, dz =$

$= -\sum_i \int \ln q_i \, q_i \, dz_i$

$= \int \ln p(x, z) \cdot \prod q_i \, dz - \int \ln \prod q_i \prod q_i \, dz$



$" \mathcal{L}(q) + \underbrace{KL(q \| p(z \mid x))}_{\min_q, \text{ i.e. } \boxed{q \approx p(z \mid x)}} = \ln p(x)$

$\int \left[ \int \ln p(x, z) \prod q_j \, dz_{-i} \right] q_i \, dz_i - \int \ln q_i \, q_i \, dz_i$

$\mathcal{L}(q) \leq \ln p(x)$

variational lower bound

$E_{z_{-i}} \ln p(x, z) = \ln \tilde{p}(x, z_i)$

$q_1^{(x_1)} q_2^{(x_2)} \approx p(x_1, x_2) =$

$\ln q_1^{*(x_1)} = E_{x_2} \ln p(x_1, x_2)$  $\begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix}^T \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix} \begin{pmatrix} \cdot \\ \cdot \end{pmatrix}$

$\max_{q_i} \leftarrow \mathcal{L}(q_i) = \text{const} + \int \ln \tilde{p} \, q_i \, dz_i - \int \ln q_i \, q_i \, dz_i$

$E_{x_2} \left[ \text{const} - \frac{1}{2} (\bar{x} - \bar{\mu})^T \Lambda (\bar{x} - \bar{\mu}) \right] = -\frac{1}{2} \left[ (x_1 - \mu_1)^T \Lambda_{11} (x_1 - \mu_1) + E_{x_2} \cancel{(x_2 - \mu_2) \Lambda_{22} (x_2 - \mu_2)} + 2 E_{x_2} (x_1 - \mu_1) \Lambda_{12} (x_2 - \mu_2) \right]$

$= \text{const} + \int \ln \frac{\tilde{p}}{q_i} q_i \, dz_i = -KL(q_i \| \tilde{p})$