Problem #1

$$\tilde{p}(x) = \exp\left(-\frac{1}{2\sigma_p^2}x^2\right), \quad p(x) = \tilde{p}(x)/Z_p$$

with

$$x^{(i)} \overset{iid}{\sim} \mathcal{N}(0,1), \quad i = 1, \dots, N$$

define

$$\hat{Z} = \frac{1}{N}\sum_{i=1}^{N} \frac{\tilde{p}(x^{(i)})}{q(x^{(i)})}$$

a) $\text{Bias}[\hat{\theta}] = \mathbb{E}[\hat{\theta}] - \theta$. Here

$$\mathbb{E}_q[\hat{Z}] = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}_q\left[\frac{\tilde{p}(x^{(i)})}{q(x^{(i)})}\right] \quad (\mathbb{E} \text{ is a linear operator})$$

$$= \frac{1}{N}\sum_{i=1}^{N} \int dx^{(i)} \frac{\tilde{p}(x^{(i)})}{q(x^{(i)})} q(x^{(i)})$$

$$= \frac{1}{N}\sum_{i=1}^{N} \int dx^{(i)} \tilde{p}(x^{(i)})$$

Note: $\int dx\, p(x) = \int dx \frac{\tilde{p}(x)}{Z_p} = 1$ ($p(x)$ normalized)

$$\Rightarrow \int dx\, \tilde{p}(x) = Z_p$$

and

$$\mathbb{E}_q[\hat{Z}] = \frac{1}{N}\sum_{i=1}^{N} Z_p = Z_p.$$

So $\text{Bias}[\hat{Z}] = Z_p - Z_p = \underline{\underline{0}}$ unbiased indeed.

b) Let $f(x) = \tilde{p}(x)/q(x)$.

Now, $\text{Var}[\hat{\theta}] = \mathbb{E}\left[(\hat{\theta} - \mathbb{E}[\hat{\theta}])^2\right]$,

so

$$\mathrm{Var}[\hat{z}] = \mathbb{E}[(\hat{z} - \mathbb{E}[\hat{z}])^2]$$

$$= \mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\tilde{p}(x^{(i)})}{q(x^{(i)})} - \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\frac{\tilde{p}(x^{(i)})}{q(x^{(i)})}\right]\right)^2\right]$$

$$= \mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^{N} f(x^{(i)}) - \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}[f(x^{(i)})]\right)^2\right]$$

$$= \mathbb{E}\left[\left(\frac{1}{N}\sum_{i=1}^{N}\left(f(x^{(i)}) - \mathbb{E}[f(x^{(i)})]\right)\right)^2\right]$$

Note: $\left(\sum_{i=1}^{N} a_i\right)^2 = (a_1 + \dots + a_N)(a_1 + \dots + a_N)$

$$= a_1a_1 + a_1a_2 + \dots + a_1a_N + \dots + a_Na_1 + \dots + a_Na_N$$

$$= a_1\left(\sum_{i=1}^{N} a_i\right) + \dots + a_N\left(\sum_{i=1}^{N} a_i\right)$$

$$= \sum_{j=1}^{N} a_j \sum_{i=1}^{N} a_i$$

$$= \sum_{i,j=1}^{N} a_i a_j$$

So $\mathbb{E}\left[\left(\sum_{i=1}^{N} a_i\right)^2\right] = \mathbb{E}\left[\sum_{i,j} a_i a_j\right] = \sum_{i,j}\mathbb{E}[a_i a_j]$

$$\mathrm{Var}[\hat{z}] = \frac{1}{N^2}\sum_{i,j=1}^{N}\mathbb{E}\left[\left(f(x^{(i)}) - \mathbb{E}[f(x^{(i)})]\right)\left(f(x^{(j)}) - \mathbb{E}[f(x^{(j)})]\right)\right]$$

inside an expectation value i & j are the same thing, so the above simplifies as

$$\frac{1}{N^2}\sum_{i=1}^{N}\underbrace{\mathbb{E}\left[\left(f(x^{(i)}) - \mathbb{E}[f(x^{(i)})]\right)^2\right]}_{=\mathrm{Var}[f(x^{(i)})]}$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{Var}[f(x^{(i)})]$$

$$= \frac{1}{N^2} N \mathrm{Var}[f(X)]$$

$$\Rightarrow \mathrm{Var}[\hat{Z}] = \mathrm{Var}[f(X)], \text{ provided } \mathrm{Var}[f(x)] < \infty$$

c) For what σ_p is $\mathrm{Var}[f(x)] < \infty$?

$$f(x) = \frac{\tilde{p}(x)}{q(x)}, \quad \tilde{p}(x) = \exp\left(-\frac{1}{2\sigma_p^2} x^2\right), \quad q(x) \sim \mathcal{N}(0,1)$$

$$\Rightarrow q(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$$

So $$f(x) = \frac{\exp\left(-\frac{1}{2\sigma_p^2} x^2\right)}{\exp\left(-\frac{1}{2}x^2\right)/\sqrt{2\pi}}$$

$$= \sqrt{2\pi} \exp\left(-\frac{1}{2\sigma_p^2} x^2 - \left(-\frac{1}{2}x^2\right)\right)$$

$$\Rightarrow f(x) = \sqrt{2\pi} \exp\left(-\frac{1}{2}x^2\left(\frac{1}{\sigma_p^2} - 1\right)\right) = \sqrt{2\pi} \exp\left(-\frac{1}{2\sigma_f^2} x^2\right)$$

$f(x)$ is an unnormalized Gaussian with variance

$$\sigma_f^2 = \left(\frac{1}{\sigma_p^2} - 1\right)^{-1}$$

which is not finite when $\frac{1}{\sigma_p^2} - 1 = 0$

$$\Rightarrow \boxed{\sigma_p^2 = 1}$$

also, if $\sigma_p^2 > 1$, then $\frac{1}{\sigma_p^2} - 1 < 0 \Rightarrow \sigma_f < 0$, not good!

So the values of σ_p^2 that keep σ_f^2 (the variance of $f(x)$) finite (and positive) are

$$\sigma_p^2 \in [0,1)$$

Problem #2

$$X_s \in [0,1], \quad p(x;\eta) = \frac{1}{Z_p}\exp\Big(\sum_{s\in V}\eta_s x_s + \sum_{\{s,t\}\in E}\eta_{st}x_s x_t\Big)$$

a) Gibbs sampling update

$$p(X_i = 1 \mid X_{\neg i}) \propto P(X_i = 1, X_{\neg i})$$

"net" i

$\neg i$ terms

$$= \exp\Big(\underbrace{\eta_i x_i}_{x_i=1} + \underbrace{\sum_{j\in N(i)}}_{\text{Neighbours of } i} \eta_{ij}\, x_i \underbrace{x_j}_{=1} + \text{rest}\Big)$$

$$= \exp\Big(\eta_i + \sum_{j\in N(i)}\eta_{ij}x_j\Big)\exp(\text{rest})$$

Renormalize to get update step:

$$p(X_i \mid X_{\neg i}) = \frac{p(X_i=1, X_{\neg i})}{\sum_{x_i} p(x_i, X_{\neg i})}$$

$$= \frac{p(x_i=1, X_{\neg i})}{p(x_i=0, X_{\neg i}) + p(x_i=1, X_{\neg i})}$$

$$= \frac{\exp\big(\eta_i + \sum_{j\in N(i)}\eta_{ij}x_j\big)\exp(\text{rest})}{\exp(0+0+\text{rest}) + \exp\big(\eta_i + \sum_{j\in N(i)}\eta_{ij}x_j\big)\exp(\text{rest})}$$

$$= \frac{\exp\big(\eta_i + \sum_{j\in N(i)}\eta_{ij}x_j\big)\exp(\text{rest})}{\big(1 + \exp\big(\eta_i + \sum_{j\in N(i)}\eta_{ij}x_j\big)\big)\exp(\text{rest})}$$

$$\Rightarrow \boxed{p(x_i=1 \mid X_{\neg i}) = \sigma\Big(u_i + \sum_{j\in N(i)} u_{ij} x_j\Big)}$$

sigmoid function.

b) Fully factorized approximation q w/ $q(X_s=1)=\tau_s$, what is

$KL(q\|p) - \log Z_p$?

$$KL(q\|p) - \log Z_p = \mathbb{E}_q\Big[\log \frac{q(x)}{p(x)}\Big] - \log Z_p$$

$$= \mathbb{E}_q[\log q(x) - \log p(x)] - \log Z_p$$

$$= \mathbb{E}_q \log q(x) \quad - \mathbb{E}_q\Big[\log\Big(\frac{1}{Z_p}\exp\Big(\sum_{s\in V} u_s X_s + \sum_{\{s,t\}\in E} u_{st} X_s X_t\Big)\Big)\Big] - \log Z_p$$

$$= \underbrace{\mathbb{E}_q \log q(x)}_{=①} \quad \cancel{\mathbb{E}_q \log Z_p} - \underbrace{\mathbb{E}_q\Big[\sum_{s\in V} u_s X_s + \sum_{\{s,t\}\in E} u_{st} X_s X_t\Big]}_{=②} - \cancel{\log Z_p}$$

$= ① + ②$

$q = \prod_s q(X_s)$ Fully factorized

$$① = \mathbb{E}_q \log q(x) = \mathbb{E}_q \log \prod_{s\in V} q(X_s) = \sum_{s\in V} \mathbb{E}_q \log q(X_s)$$

using $q(X_s=1)=\tau_s \Rightarrow q(X_s=0) = 1 - q(X_s=1) = 1-\tau_s$

$$= \sum_{s\in V} \sum_{X_s=0,1} q(X_s)\log q(X_s)$$

$$\Rightarrow ① = \sum_{s\in V}\big((1-\tau_s)\log(1-\tau_s) + \tau_s \log \tau_s\big)$$

$$② = \mathbb{E}_q\Big[\sum_{s\in V} u_s X_s + \sum_{\{s,t\}\in E} u_{st} X_s X_t\Big]$$

$$= \sum_{s\in V} u_s \mathbb{E}_q[X_s] + \sum_{\{s,t\}\in E} u_{st}\, \mathbb{E}_q[X_s X_t]$$

$$= \sum_{s\in V} \sum_{x_s=0,1} \eta_s\, q(x_s) x_s + \sum_{\{s,t\}\in E} \eta_{st}\, \mathbb{E}_q[x_s]\mathbb{E}_q[x_t]$$

because $q(x) = \prod_s q(x_s)$

x_s and x_t are independent and

$\mathbb{E}_q(x_s x_t) = \mathbb{E}_q[x_s]\mathbb{E}_q[x_t]$

$$\Rightarrow ② = \sum_{s\in V} \eta_s(\tau_s(1) + (1-\tau_s)(0))$$

$$+ \sum_{\{s,t\}\in E} \eta_{st} \sum_{x_s=0,1} \sum_{x_t=0,1} q(x_s) x_s\, q(x_t)$$

$$= \sum_{s\in V} \eta_s \tau_s + \sum_{\{s,t\}\in E} \eta_{st} (\tau_s(1)\tau_t(1) + \tau_s(1)(1-\tau_t)(0) + (1-\tau_s)(0)$$

$\tau_t(1) + (1-\tau_s)(0)(1-\tau_t)(0)$.

$$\Rightarrow ② = \sum_{s\in V} \eta_s \tau_s + \sum_{\{s,t\}\in E} \eta_{st} \tau_s \tau_t$$

and

$$\boxed{\begin{aligned} KL(q\|p) - \log Z_p = &\sum_{s\in V}(\tau_s \log \tau_s + (1-\tau_s)\log(1-\tau_s)) \\ &- \sum_{s\in V} \eta_s\tau_s - \sum_{\{s,t\}\in E} \eta_{st}\tau_s\tau_t \end{aligned}}$$

c) Find update rule for τ_s by minimizing $KL(q\|p)$.
Since Z_p does not depend on τ_s, I can minimize $KL(q\|p) - \log Z_p$ with respect to τ_s:

$$\frac{\partial}{\partial \tau_s}(KL(q\|p) - \log Z_p) = 0$$

$$\Rightarrow \frac{\partial}{\partial \tau_s}\Big(\sum_{k\in V}\big(\tau_k \log \tau_k + (1-\tau_k)\log(1-\tau_k)\big) - \sum_{k\in V} \eta_k \tau_k - \sum_{\{k,t\}\in E} \eta_{kt}\tau_k\tau_t \Big) = 0$$

$$\Rightarrow \log \tau_s + \frac{\cancel{\tau_s}}{\cancel{\tau_s}} - \log(1-\tau_s) + \frac{\cancel{(1-\tau_s)}}{-\cancel{(1-\tau_s)}} - \eta_s - \sum_{t\in N(s)} \eta_{st}\tau_t = 0$$

$$\Rightarrow \log \tau_s - \log(1-\tau_s) = \eta_s + \sum_{t\in N(s)} \eta_{st}\tau_t$$

$$\Rightarrow \log\left(\frac{\tau_s}{1-\tau_s}\right) = \eta_s + \sum_{t\in N(s)} \eta_{st}\tau_t$$

$$\Rightarrow \frac{\tau_s}{1-\tau_s} = \exp\Big(\eta_s + \sum_{t\in N(s)} \eta_{st}\tau_t\Big)$$

$$\Rightarrow \tau_s = (1-\tau_s)\exp\Big(\eta_s + \sum_{t\in N(s)} \eta_{st}\tau_t\Big)$$

$$\Rightarrow \tau_s\Big(1+\exp\Big(\eta_s + \sum_{t\in N(s)} \eta_{st}\tau_t\Big)\Big) = \exp\Big(\eta_s + \sum_{t\in N(s)} \eta_{st}\tau_t\Big)$$

$$\Rightarrow \tau_s = \frac{\exp\Big(\eta_s + \sum_{t\in N(s)} \eta_{st}\tau_t\Big)}{1 + \exp\Big(\eta_s + \sum_{t\in N(s)} \eta_{st}\tau_t\Big)}$$

$$\Rightarrow \boxed{\tau_s = \sigma\Big(\eta_s + \sum_{t\in N(s)} \eta_{st}\tau_t\Big)} \quad (*)$$

Sigmoid function

Mean field updates: update τ_s according to $(*)$