LATENT DISCRETE GRAPHICAL MODEL

Measurement M, latent state L, parameters $\Theta = (\delta, \gamma, \theta, \rho, D)$, case/control indicator Y.

Bronze standard only $[M, L, \Theta] = \{\prod_{k=1}^{K}[M_k|L_k, \Theta]\}[L|\Theta][\Theta]$

$$[M, L, \Theta] \propto \exp\Big\{\sum_{k=1}^{K}\Big[\log(1-\delta_k) + M_k \log\frac{\delta_k}{1-\delta_k} + YL_k \log\frac{1-\gamma_k}{1-\delta_k} + YL_k M_k \log\frac{\gamma_k(1-\delta_k)}{\delta_k(1-\gamma_k)}\Big] +$$
$$Y\Big[\sum_{k=1}^{K}(L_k\theta_k + \rho\sum_{k\neq k'} L_k L_{k'} D_{kk'}) - A(\theta, \rho, D)\Big]\Big\}[\Theta]$$

Variational approximate posterior distribution of Θ.

Assume the approximate posterior $Q(\Theta, L)$ to be fully factorized, i.e.

$$Q(\Theta, L) = q(\delta, \gamma)q(L)q(\theta)q(\rho)q(D) = \prod_{\eta\in\{\Theta, L\}} q(\eta)$$

We maximize the lower bound of the joint density (minimize the KL divergence) to find Q.

$$Q(\Theta, L) = \mathrm{argmax}\{\mathbb{E}_Q\Big[\log\frac{[M, L, \Theta]}{Q(\Theta, L)}\Big]\}$$
$$q(\eta) \propto \exp\{\mathbb{E}_{Q(-\eta)}(\log[M, L, \Theta])\}$$

For q(γ, δ):

$$
\begin{aligned}
q(\delta,\gamma) &\propto \exp\Big\{\mathbb{E}_L\big[\sum_{k=1}^{K}\log[M_k|L_k,\delta,\gamma]+\log[\delta,\gamma]\big]\Big\} \\
&\propto \exp\Big\{\sum_{k=1}^{K}\Big[\log(1-\delta_k)+M_k\log\frac{\delta_k}{1-\delta_k}+Yq_{l_k}\big(\log\frac{1-\gamma_k}{1-\delta_k}\big)+Yq_{l_k}M_k\log\frac{\gamma_k(1-\delta_k)}{\delta_k(1-\gamma_k)}\Big]\Big\}[\gamma,\delta] \\
&= \prod_{k=1}^{K}\Big[(1-\gamma_k)^{Yq_{l_k}(1-M_k)}\gamma_k^{Yq_{l_k}M_k}\Big]\Big[(1-\delta_k)^{1+Yq_{l_k}M_k-Yq_{l_k}-M_k}\delta_k^{M_k(1-Yq_{l_k})}\Big][\gamma,\delta]
\end{aligned}
$$

Where $q_{l_k}=\mathbb{E}_Q(L_k)$. Let a_k^*, b_k^* be the hyper-parametes for γ_k, and let c_k^*, d_k^* be the hyper-parameters for δ_k. Then the posterior distribution is: $\gamma_k \sim \mathrm{Beta}(A_k^*, B_k^*)$ and $\delta_k \sim \mathrm{Beta}(A_k', B_k')$ where

$$
\begin{aligned}
A_k^* &= Yq_{l_k}M_k + a_k^* \\
B_k^* &= Yq_{l_k}(1-M_k) + b_k^* \\
A_K' &= M_k(1-Yq_{l_k}) + c_K^* \\
B_K' &= 1 + Yq_{l_k}M_k - Yq_{l_k} - M_k + d_K^*
\end{aligned}
$$

With n independent samples:

$$
\begin{aligned}
A_{kn}^* &= \sum_{i=1}^{n} Y_i q_{l_{ik}} M_{ik} + a_k^* \\
B_{kn}^* &= \sum_{i=1}^{n} Y_i q_{l_{ik}}(1-M_{ik}) + b_k^* \\
A_{kn}' &= \sum_{i=1}^{n} M_{ik}(1-Y_i q_{l_{ik}}) + c_K^* \\
B_{kn}' &= n + \sum_{i=1}^{n}(Y_i q_{l_{ik}} M_{ik} - Y_i q_{l_{ik}} - M_{ik}) + d_K^*
\end{aligned}
$$

For $q(L)$:

$$
\begin{aligned}
q(L) = \prod_k q(L_k) &\propto \exp\Big\{\mathbb{E}_\Theta\big[\sum_{k=1}^{K}\log[M_k|L_k,\delta,\gamma] + \log[L|\theta,\rho,D]\big]\Big\} \\
&= \exp\Big\{Y\mathbb{E}\sum_{k=1}^{K}\Big[L_k(\theta_k + \log\frac{1-\gamma_k}{1-\delta_k} + M_k\log\frac{\gamma_k(1-\delta_k)}{\delta_k(1-\gamma_k)}) + \rho\sum_{k'\neq k} L_k L_{k'} D_{kk'}\Big]\Big\}
\end{aligned}
$$

Let $h_k = \theta_k + \log\frac{1-\gamma_k}{1-\delta_k} + M_k\log\frac{\gamma_k(1-\delta_k)}{\delta_k(1-\gamma_k)}$, and let $d_{kk'} = \mathbb{E}_Q(D_{kk'})$, then

$$
\begin{aligned}
q(L_k) &\propto \exp\Big\{Y\Big[L_k\mathbb{E}(h_k) + \rho L_k\sum_{k'\neq k} q_{l_{k'}} d_{kk'}\Big]\Big\} \\
q_{l_k} &= \frac{\exp[Y(H_k + \mu_\rho\sum_{k'\neq k} q_{l_{k'}} d_{kk'})]}{1+\exp[Y(H_k + \mu_\rho\sum_{k'\neq k} q_{l_{k'}} d_{kk'})]}
\end{aligned}
$$

Where with $\psi(x) = \frac{\partial\log\Gamma(x)}{\partial x}$ and $\psi(x+y) = \frac{\partial\log\Gamma(x+y)}{\partial x}$

$$
\begin{aligned}
H_k = \mathbb{E}_Q(h_k) = &M_k[\psi(A_k^*) - \psi(A_k')] + (1-M_k)[\psi(B_k^*) - \psi(B_k')] \\
&+ \mu_{\theta_k} - \psi(A_k^* + B_k^*) + \psi(A_k' + B_k')
\end{aligned}
$$

with n independent samples, indexed by i:

$$
\begin{aligned}
q_{l_{ik}} &= \frac{\exp[Y_i(H_{ik} + \mu_\rho\sum_{k'\neq k} q_{l_{ik'}} d_{kk'})]}{1+\exp[Y_i(H_{ik} + \mu_\rho\sum_{k'\neq k} q_{l_{ik'}} d_{kk'})]} \\
H_{ik} &= M_{ik}[\psi(A_{kn}^*) - \psi(A_{kn}')] + (1-M_{ik})[\psi(B_{kn}^*) - \psi(B_{kn}')] \\
&\quad + \mu_{\theta_k} - \psi(A_{kn}^* + B_{kn}^*) + \psi(A_{kn}' + B_{kn}')
\end{aligned}
$$

For $q(\theta)$:

$$q(\theta) = \prod_k q(\theta_k) \propto \exp\Big\{Y\mathbb{E}_{(L,\rho,D)}\big[\sum_{k=1}^{K} L_k\theta_k - A(\theta,\rho,D)\big]\Big\}[\theta]$$

$$\approx \exp\Big\{Y\sum_{k=1}^{K}\big[q_{l_k}\theta_k - \mathbb{E}\log\big(1+\exp(\theta_k+\rho\sum_{k'\neq k} L_{k'}D_{kk'})\big)\big]\Big\}[\theta] \text{ by pseudo-likelihood}$$

By taylor expansion,

$$\mathbb{E}_y \log(1+e^{x+y}) = \log(1+e^{x+\mu_y}) + \frac{e^{x+\mu_y}}{2(1+e^{x+\mu_y})^2}\mathbb{V}\text{ar}(y) + \mathbb{E}_y[o(y-\mu_y)^3]$$

Let $R_k = \sum_{k'\neq k} L_{k'}D_{kk'}$, then

$$\begin{aligned}
\mathbb{E}(R_k) &= \sum_{k'\neq k} q_{l_{k'}} d_{kk'} \\
\mathbb{V}\text{ar}(R_k) &= \sum_{k'\neq k} q_{l_{k'}} d_{kk'}(1-q_{l_{k'}} d_{kk'}) \\
\mathbb{V}\text{ar}(\rho R_k) &= \mathbb{E}(\rho^2)\mathbb{E}(R_k^2) - \mathbb{E}(\rho)^2\mathbb{E}(R_k)^2 \\
&= (\sigma_\rho^2+\mu_\rho^2)[\mathbb{V}\text{ar}(R_k)+\mathbb{E}(R_k)^2] - \mu_\rho^2\mathbb{E}(R_k)^2 \\
&= (\sigma_\rho^2+\mu_\rho^2)\sum_{k'\neq k} q_{l_{k'}} d_{kk'}(1-q_{l_{k'}} d_{kk'}) + \sigma_\rho^2[\sum_{k'\neq k} q_{l_{k'}} d_{kk'}]^2
\end{aligned}$$

Thus, $q(\theta_k)$ is further approximated by

$$q(\theta_k) \propto \exp\Big\{Y\big[q_{l_k}\theta_k - \log(1+e^{\theta_k+\mu_\rho\sum_{k'\neq k} q_{l_{k'}} d_{kk'}}) - \frac{e^{\theta_k+\mu_\rho\sum_{k'\neq k} q_{l_{k'}} d_{kk'}}}{2(1+e^{\theta_k+\mu_\rho\sum_{k'\neq k} q_{l_{k'}} d_{kk'}})^2}\mathbb{V}\text{ar}(\rho R_k)\big] - \frac{1}{2}(\theta_k-\mu_\theta^*)^2\tau_\theta^*\Big\}$$

with n independent samples:

$$q(\theta_k) \propto \exp\Big\{\sum_{i=1}^{n} Y_i\Big[q_{l_{ik}}\theta_k - \log(1+e^{\theta_k+\mu_\rho\sum_{k'\neq k} q_{l_{ik'}} d_{kk'}}) - \frac{e^{\theta_k+\mu_\rho\sum_{k'\neq k} q_{l_{ik'}} d_{kk'}}}{2(1+e^{\theta_k+\mu_\rho\sum_{k'\neq k} q_{l_{ik'}} d_{kk'}})^2}\mathbb{V}\text{ar}(\rho R_{ik})\Big] - \frac{1}{2}(\theta_k-\mu_\theta^*)^2\tau_\theta^*\Big\}$$

$$\mathbb{V}\text{ar}(\rho R_{ik}) = (\sigma_\rho^2+\mu_\rho^2)\sum_{k'\neq k} q_{l_{ik'}} d_{kk'}(1-q_{l_{ik'}} d_{kk'}) + \sigma_\rho^2[\sum_{k'\neq k} q_{l_{ik'}} d_{kk'}]^2$$

Then $\mu_{\theta_k} \approx \mathbf{argmax}\ q(\theta_k)$, and $\sigma^2_{\theta_k} \approx -\Big[\frac{\partial^2 \log q(\theta_k)}{\partial \theta_k^2}\Big|_{\theta_k=\mu_{\theta_k}}\Big]^{-1}$.

For $q(\rho)$:

$$q(\rho) \propto \exp\Big\{Y\mathbb{E}_{(\theta,L,D)}\,[\rho\sum_{k=1}^{K}\sum_{k'\neq k} L_k L_{k'} D_{kk'} - A(\theta,\rho,D)]\Big\}[\rho]$$

$$\approx \exp\Big\{Y\rho\sum_{k=1}^{K}[\sum_{k'\neq k} q_{l_k} q_{l_{k'}} d_{kk'} - \mathbb{E}\log\big(1+\exp(\theta_k+\rho\sum_{k'\neq k} L_{k'} D_{kk'})\big)]\Big\}[\rho]$$

By Le Cam's theorem, R_k can be approximated by Poisson$(\sum_{k'\neq k} q_{l_{k'}} d_{kk'})$, thus

$$q(\rho) \propto \exp\Big\{Y\rho\sum_{k=1}^{K}[\sum_{k'\neq k} q_{l_k} q_{l_{k'}} d_{kk'} - \sum_{j=0}^{K-1}\text{Poi}(j;\lambda=\sum_{k'\neq k} q_{l_{k'}} d_{kk'})\Big(\log(1+e^{\mu_{\theta_k}+j\rho}) + \frac{e^{\mu_{\theta_k}+j\rho}}{2(1+e^{\mu_{\theta_k}+j\rho})^2}\sigma^2_{\theta_k}\Big)]\Big\}[\rho]$$

With n independent samples:

$$q(\rho) \propto \exp\Big\{\rho\sum_{i=1}^{n} Y_i\sum_{k=1}^{K}[\sum_{k'\neq k} q_{l_{ik}} q_{l_{ik'}} d_{kk'} - \sum_{j=0}^{K-1}\text{Poi}(j;\lambda=\sum_{k'\neq k} q_{l_{ik'}} d_{kk'})\Big(\log(1+e^{\mu_{\theta_k}+j\rho}) + \frac{e^{\mu_{\theta_k}+j\rho}}{2(1+e^{\mu_{\theta_k}+j\rho})^2}\sigma^2_{\theta_k}\Big)]\Big\}[\rho]$$

Then $\mu_\rho \approx \mathbf{argmax}\ q(\rho)$, and $\sigma^2_\rho \approx -\Big[\frac{\partial^2 \log q(\rho)}{\partial \rho^2}\Big|_{\rho=\mu_\rho}\Big]^{-1}$.

For $q(D)$:

$$q(D) \propto \exp\Big\{Y\mathbb{E}_{(\theta,L,\rho)}\,[\rho\sum_{k=1}^{K}\sum_{k'\neq k} L_k L_{k'} D_{kk'} - A(\theta,\rho,D)]\Big\}[D]$$

$$\approx \exp\Big\{Y\rho\sum_{k=1}^{K}[\sum_{k'\neq k} q_{l_k} q_{l_{k'}} d_{kk'} - \mathbb{E}\log\big(1+\exp(\theta_k+\rho\sum_{k'\neq k} L_{k'} D_{kk'})\big)]\Big\}[D]$$

$$q(D_{k_1k_2}) \propto \exp\Big\{2Y\mu_\rho q_{l_{k_1}} q_{l_{k_2}} D_{k_1k_2} - Y\mathbb{E}\log(1+e^{\theta_{k_1}+\rho\sum_{k\neq k_1,k\neq k_2} L_k D_{kk_1}+\rho L_{k_2} D_{k_1k_2}})$$
$$- Y\mathbb{E}\log(1+e^{\theta_{k_2}+\rho\sum_{k\neq k_1,k\neq k_2} L_k D_{kk_2}+\rho L_{k_1} D_{k_1k_2}}) + D_{k_1k_2}\mathbb{E}_{\pi_d}(\log\frac{\pi_d}{1-\pi_d})\Big\}$$

Let $S_{k_1k_2} = \theta_{k_1} + \rho\sum_{k\neq k_1,k\neq k_2} L_k D_{kk_1}$, $T_{k_1k_2} = \mathbb{E}\log(1+e^{S_{k_1k_2}+\rho L_{k_2}})$, $T'_{k_1k_2} = \mathbb{E}\log(1+e^{S_{k_1k_2}})$, and let A_{π_d}, B_{π_d} be the Beta hyper-parameters for π_d, then for all $k_1 < k_2$:

$$\mathbb{E}_{\pi_d}(\log\frac{\pi_d}{1-\pi_d}) = \psi(A_{\pi_d}) - \psi(B_{\pi_d})$$

$$\mathbb{V}\mathrm{ar}(S_{k_1k_2}) = \sigma^2_{\theta_{k_1}} + (\sigma^2_\rho + \mu^2_\rho)\sum_{k\neq k_1,k\neq k_2} q_{l_k} d_{kk_1}(1 - q_{l_k} d_{kk_1}) + \sigma^2_\rho[\sum_{k\neq k_1,k\neq k_2} q_{l_k} d_{kk_1}]^2$$

$$\mathbb{V}\mathrm{ar}(\rho L_k) = (\sigma^2_\rho + \mu^2_\rho)q_{l_k} - \mu^2_\rho q^2_{l_k}$$

Thus

$$T_{k_1k_2} \approx \log(1+e^{\mu_{\theta_{k_1}}+\mu_\rho\sum_{k\neq k_1,k\neq k_2} q_{l_k}d_{kk_1}+\mu_\rho q_{l_{k_2}}}) + \frac{e^{\mu_{\theta_{k_1}}+\mu_\rho\sum_{k\neq k_1,k\neq k_2} q_{l_k}d_{kk_1}+\mu_\rho q_{l_{k_2}}}}{2(1+e^{\mu_{\theta_{k_1}}+\mu_\rho\sum_{k\neq k_1,k\neq k_2} q_{l_k}d_{kk_1}+\mu_\rho q_{l_{k_2}}})^2}[\mathbb{V}\mathrm{ar}(S_{k_1k_2}) + \mathbb{V}\mathrm{ar}(\rho L_{k_2})]$$

$$T'_{k_1k_2} \approx \log(1+e^{\mu_{\theta_{k_1}}+\mu_\rho\sum_{k\neq k_1,k\neq k_2} q_{l_k}d_{kk_1}}) + \frac{e^{\mu_{\theta_{k_1}}+\mu_\rho\sum_{k\neq k_1,k\neq k_2} q_{l_k}d_{kk_1}}}{2(1+e^{\mu_{\theta_{k_1}}+\mu_\rho\sum_{k\neq k_1,k\neq k_2} q_{l_k}d_{kk_1}})^2}\mathbb{V}\mathrm{ar}(S_{k_1k_2})$$

$$d_{k_1k_2} = \frac{\exp\{2Y\mu_\rho q_{l_{k_1}} q_{l_{k_2}} - YT_{k_1k_2} - YT_{k_2k_1} + \psi(A_{\pi_d}) - \psi(B_{\pi_d})\}}{\exp\{2Y\mu_\rho q_{l_{k_1}} q_{l_{k_2}} - YT_{k_1k_2} - YT_{k_2k_1} + \psi(A_{\pi_d}) - \psi(B_{\pi_d})\} + \exp\{-YT'_{k_1k_2} - YT'_{k_2k_1}\}}$$

With n independent samples:

$$d_{k_1k_2} = \frac{\exp\{\sum_{i=1}^n Y_i[2\mu_\rho q_{l_{ik_1}} q_{l_{ik_2}} - T^{(i)}_{k_1k_2} - T^{(i)}_{k_2k_1}] + \psi(A_{\pi_d}) - \psi(B_{\pi_d})\}}{\exp\{\sum_{i=1}^n Y_i[2\mu_\rho q_{l_{ik_1}} q_{l_{ik_2}} - T^{(i)}_{k_1k_2} - T^{(i)}_{k_2k_1}] + \psi(A_{\pi_d}) - \psi(B_{\pi_d})\} + \exp\{-\sum_{i=1}^n Y_i[T'^{(i)}_{k_1k_2} + T'^{(i)}_{k_2k_1}]\}}$$