CSC412 / CSC2506 Solutions to Sample Problems for Midterm

1. Let $p(k)$ be a one-dimensional discrete distribution that we wish to approximate, with support on non-negative integers. One way to fit an approximating distribution $q(k)$ is to minimize the Kullback-Leibler divergence:

$$KL(p||q) = \sum_{k=0}^{\infty} p(k) log \frac{p(k)}{q(k)}$$

   Show that when $q(k)$ is a Poisson distribution,

$$q(k) = \frac{\lambda^k \exp(-\lambda)}{k!}$$

   this KL divergence is minimized by setting λ to the mean of $p(k)$.

   Solution:

$$\frac{\partial KL}{\partial \lambda} = 0 \implies \lambda = E[p(k)]$$

2. Recall that the definition of an exponential family model is:

$$f(x|\eta) = h(x)g(\eta)\exp(\eta^\top T(x))$$

   where:

$$\begin{aligned} \eta &\quad \text{are the parameters} \\ T(x) &\quad \text{are the sufficient statistics} \\ h(x) &\quad \text{is the base measure} \\ g(\eta) &\quad \text{is the normalizing constant} \end{aligned}$$

   Consider the univariate Gaussian, with mean μ and precision $\lambda = \frac{1}{\sigma^2}$:

$$p(D|\mu,\lambda) = \prod_{i=1}^{N} (\frac{\lambda}{2\pi})^{\frac{1}{2}} \exp(-\frac{1}{2}(x_i - \mu)^2)$$

   What are η and $T(x)$ for this distribution when represented in exponential family form?

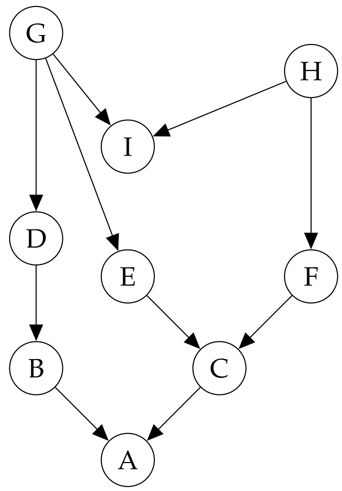
   Solution:

$$p(D|\mu,\lambda) = (2\pi)^{-N/2}[\lambda^{1/2}\exp(-\frac{\lambda}{2}\mu^2)]^N \exp[\mu\lambda \sum_i x_i - \lambda/2 \sum_i x_i^2]$$

$$\eta = \begin{bmatrix} \mu\lambda \\ -\lambda/2 \end{bmatrix}$$

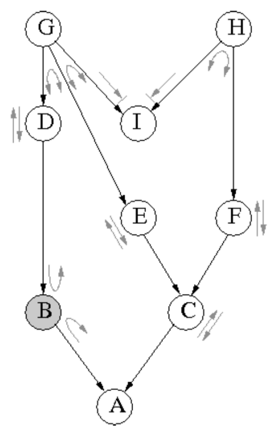
$$T(x) = \begin{bmatrix} \sum_i x_i \\ \sum_i x_i^2 \end{bmatrix}$$

3. Consider the following directed graphical model:



(a) List all variables that are independent of A given evidence on B.
(b) Write down the factorized normalized joint distribution that this graphical model represents.

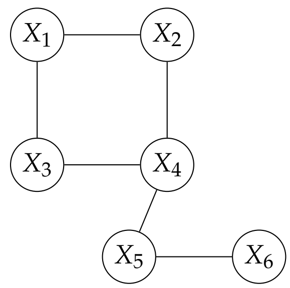
Solution: a) $\{\}$



b) $P(G)P(H)P(I|G,H)P(D|G)P(E|G)P(F|H)P(B|D)P(C|E,F)P(A|B,C)$

4. Murphy 20.1

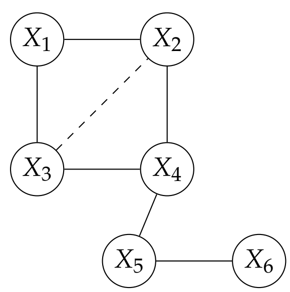
Correction! Older copies of the Murphy book have a typo pointing you to an incorrect figure. Do this question but with this MRF:



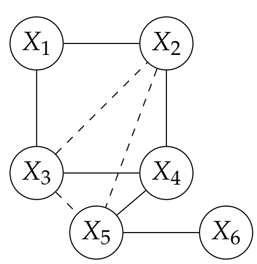
Solution:

a) Largest intermediate term has size 3 (1,2,3) and (2,3,4)

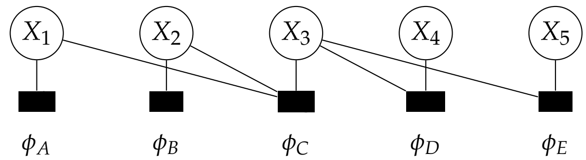
b) Largest maximal clique has size 3.



c) The largest intermediate term has size 4 (2,3,4,5)
d) Largest maximal clique has size 4.



5. Consider the Factor Graph:



(a) Write down the normalized joint distribution $P(X_1, X_2, X_3, X_4, X_5)$ in terms of the potentials.
(b) Write down any conditional independence relationships given by the graph.

Solution:

a) $\frac{1}{Z}\phi_A(X_1)\phi_B(X_2)\phi_C(X_1, X_2, X_3)\phi_D(X_3, X_4)\phi_E(X_5)$
where $Z = \sum_{\mathbf{x}} \phi_A(X_1)\phi_B(X_2)\phi_C(X_1, X_2, X_3)\phi_D(X_3, X_4)\phi_E(X_3, X_5)$

b) $X_1, X_2 \perp X_4, X_5 | X_3$
$X_4 \perp X_5 | X_3$