سوال ۱

الف) نادرست، با استفاده از الگوریتم PC ممکن است نتوانیم جهت تمام یال ها را بیابیم.

ب) درست، وقتی توزیع مقصد را پنهان به سمت یک توزیع احتمال مشخص میل می کند، باعث می شود
که Decoder یاد بگیرد که چگونه نمونه های آن توزیع احتمال را بازسازی کند و این عامل
باعث معنا دار شدن فضای پنهان می شود.

ج) درست، زیرا سرویس A و H توسط D بلاک می شود.

د) درست، اگر G، given شود، ارتباط H با کل گراف قطع می شود. لذا
متغیر H به شرط G از سایر متغیر ها مستقل است.

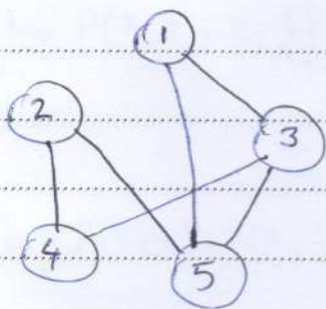ه) نادرست، چون F به صورت غیر مستقیم معلول (effect) متغیر A محسوب
می شود، داریم:

$$P(A \mid do \, F = f) = P(A)$$

رابطه ی بالا به علت اصل مکانیزم های مستقل برقرار است.

و) درست، زیرا متغیر C علت متغیر F است.

الف) از آن جایی که فاکتورها روی کلیک‌ها تعریف می‌شوند، داریم:

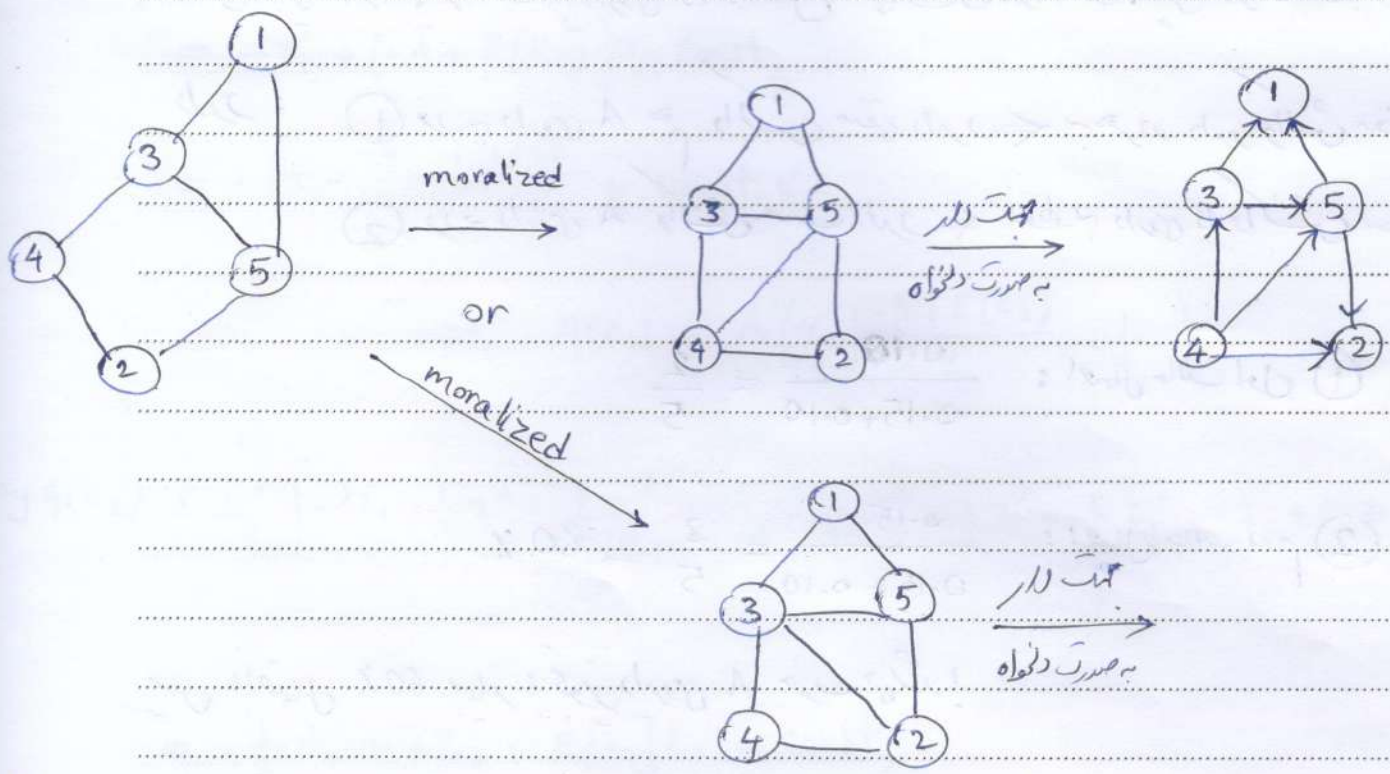

ب)

هیچ تغییری نمی‌کند، چون اگر ϕ_2، ۵ برابر شود، ضریب نرمال کننده Z نیز باید ۵ برابر گردد.

ج)

گراف بدون جهت را با چینش بهتر دوباره رسم می‌کنیم:

گراف جهت‌دار بدون جهتی که ماختار آن مشابه یکی از دو گراف moralized بالا باشد، درست است.

د) با استفاده از Rejection Sampling میتوان این کار را انجام داد. با ید یک توزیع Q

را بیابیم به طوری که:

$$kQ(x) \geqslant \tilde{P}(x) \qquad \forall x$$

$$\tilde{P}(x) = \phi_1 \phi_2 \phi_3 \phi_4$$

سوال ۳

از آن جایی که فرد بیمار نسبت به داروی B واکنش منفی نشان داده است، پس دو حالت وجود

دارد: ① فرد به داروی A هم واکنش منفی دارد ⇐ به هر دو دارو واکنش منفی دارد.

② فرد به داروی A واکنش مثبت دارد ⇐ فقط به داروی B واکنش منفی دارد.

① احتمال حالت اول: $\dfrac{0.10}{0.15 + 0.10} = \dfrac{2}{5}$

② احتمال حالت دوم: $\dfrac{0.15}{0.15 + 0.10} = \dfrac{3}{5} = 60\%$

پس با احتمال 60% بیمار با گزینه داروی A خوب میشود!

$$\log q(z_1) = E_{z_2}\left[\log P(x, z_1, z_2)\right] + const$$

$$\log q(z_2) = E_{z_1}\left[\log P(x, z_1, z_2)\right] + const$$

$$\log P(x, z_1, z_2) = \log P(z_1)\, P(x|z_1)\, P(z_2|x, z_1)$$

$$= \log \lambda - \lambda z_1 - \log \sqrt{2\pi} - \frac{1}{2}(x - z_1)^2 - \log\sqrt{2\pi} - \frac{1}{2}(z_2 - x - z_1)^2$$

$$\log q(z_1) = E_{z_2}\left[-\lambda z_1 - \overbrace{\frac{1}{2}x^2}^{const} - \frac{1}{2}z_1^2 + xz_1 - \frac{1}{2}z_2^2 - \overbrace{\frac{1}{2}x^2}^{const} - \frac{1}{2}z_1^2 + xz_2 + z_1 z_2\right.$$

$$\left. - xz_1\right] + const$$

$$= -\lambda z_1 - \frac{1}{2}z_1^2 + xz_1 - \frac{1}{2}z_1^2 + z_1\, E[z_2] - xz_1 + const_2$$

$$= -z_1^2 + (-\lambda + E[z_2])z_1 + const_2$$

$$= -\left(z_1 - \frac{-\lambda + E[z_2]}{2}\right)^2 + const_3$$

$$\Rightarrow q(z_1) = N\left(z_1 \,\middle|\, \frac{-\lambda + E[z_2]}{2},\ \frac{1}{2}\right)$$

$$\log q(z_2) = E_{z_1}\left[\underbrace{-\lambda z_1 - \frac{1}{2}x^2 - \frac{1}{2}z_1^2 + xz_1}_{Const} - \frac{1}{2}z_2^2 \underbrace{- \frac{1}{2}x^2 - \frac{1}{2}z_1^2}_{const} + xz_2 + z_1 z_2 \overset{const}{- xz_1}\right]$$
$$+ const$$

$$= -\frac{1}{2}z_2^2 + xz_2 + E[z_1]z_2 + const_2$$

$$= -\frac{1}{2}\left(z_2^2 - 2(x + E[z_1])z_2\right) + const_2 = -\frac{1}{2}\left(z_2 - (x + E[z_1])\right)^2 + const_3$$

$$\implies q(z_2) = N\left(z_2 \mid x + E[z_1], 1\right)$$

حق با احمد است. چپ بار (CD) به صورت $C \to D$ خواهد بود. زیرا در غیر
این صورت با یک v-structure جدید ایجاد می‌شود و با (CD) در ایجاد می‌شود.

در حالت عالی بایستی نشان دهیم که از جهت به صورت $C \to D$ باشد یکی از دو شکل
بالا برخورد می‌کنیم.

$$\min_{\alpha} KL\left(q(z) \,\|\, p(x,z)\right)$$

$$KL\left(q(z) \,\|\, p(x,z)\right) = E_q\left[\log \frac{q(z)}{p(x,z)}\right] = E_q\left[\log \frac{q(z)}{p(z)\,p(x|z)}\right]$$

$$= E_q\left[\log \frac{\frac{z}{\alpha} e^{-\frac{z^2}{2\alpha}}}{z\, e^{-\frac{z^2}{2}} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{(x-z)^2}{-2}\right)}\right]$$

$$= E_q\left[-\log \alpha - \frac{z^2}{2\alpha} + \frac{z^2}{2} + \frac{(x-z)^2}{2} + \log \sqrt{2\pi}\right]$$

$$= -\log \alpha - \frac{E[z^2]}{2\alpha} + \frac{E[z^2]}{2} + \frac{1}{2}\left(x^2 - 2x E[z] + E[z^2]\right) + \log \sqrt{2\pi}$$

$$= -\log \alpha + \underbrace{E[z^2]}_{2\alpha} \left(\frac{-1}{2\alpha} + 1\right) - x E[z] + const$$

$$\underbrace{\qquad}_{\sqrt{\frac{\alpha \pi}{2}}}$$

$$= -\log \alpha - 1 + 2\alpha - x\sqrt{\frac{\alpha\pi}{2}} + const$$

$$\frac{d\,KL}{d\alpha} = \frac{-1}{\alpha} + 2 - x\sqrt{\frac{\pi}{2}}\,\frac{1}{2}\,\frac{1}{\sqrt{\alpha}} = 0$$

ضرب در α ها طرفین $\Longrightarrow$ $-1 + 2\alpha - \sqrt{\alpha}\,\underbrace{\frac{x}{2}\sqrt{\frac{\pi}{2}}}_{a} = 0$

می‌کنیم

$\sqrt{\alpha} = t \Longrightarrow$ $\quad 2t^2 - ta - 1 = 0$

$$t = \frac{a \pm \sqrt{a^2 + 8}}{4}$$

چون $t > 0$ بنابراین؛

$$t = \frac{a + \sqrt{a^2 + 8}}{4} \qquad \sout{\alpha = \frac{\sqrt{a + \sqrt{}}}{2}}$$

$$\Longrightarrow \alpha = \left(\frac{a \pm \sqrt{a^2 + 8}}{4}\right)^2 \quad , \quad a = \frac{x}{2}\sqrt{\frac{\pi}{2}}$$

الف) بله، زیرا این تبدیل معکوس‌پذیر است:

$$x_i = z_i \, e^{\alpha_i} + \mu_i \quad \Longleftrightarrow \quad z_i = \frac{x_i - \mu_i}{e^{\alpha_i}}$$

ب)

$$J = \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{array} \begin{bmatrix} \dfrac{\partial x_1}{\partial z_1} & \dfrac{\partial x_1}{\partial z_2} & \cdots & \dfrac{\partial x_1}{\partial z_n} \\[2mm] \dfrac{\partial x_2}{\partial z_1} & \dfrac{\partial x_2}{\partial z_2} & \cdots & \dfrac{\partial x_2}{\partial z_n} \\[2mm] \vdots & & \ddots & \\[2mm] \dfrac{\partial x_n}{\partial z_1} & \dfrac{\partial x_n}{\partial z_2} & \cdots & \dfrac{\partial x_n}{\partial z_n} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ ? & e^{\alpha_2} & 0 & \cdots & 0 \\ ? & ? & e^{\alpha_3} & 0 & \cdots \\ ? & ? & \cdots & \ddots & \\ ? & ? & \cdots & & e^{\alpha_n} \end{bmatrix}$$

چون ماتریس پایین مثلثی است، دترمینان آن حاصل‌ضرب درایه‌های قطر اصلی:

$$|J| = 1 \times e^{\alpha_2} \times e^{\alpha_3} \times \cdots \times e^{\alpha_n} = e^{\alpha_2 + \cdots + \alpha_n}$$

ج) حالا اینکه شبکه‌های عصبی α_i و μ_i را روی $[x_1, \ldots, x_{i-1}]$ و سپس اعمال کنیم آن‌ها را روی $[z_1, \ldots, z_{i-1}]$

احتمال کنیم. در این صورت، همهٔ z_i‌ها را ابتدا تولید می‌کنیم و به صورت موازی x_i‌ها را می‌سازیم. در این حالت برای محاسبهٔ x_i به x_1 تا x_{i-1} نیازی نداریم.