

توجه: نمره امتحان از ۱۱۰ است و برای کامل شدن باید ۱۰۰ نمره را کسب نمایید.

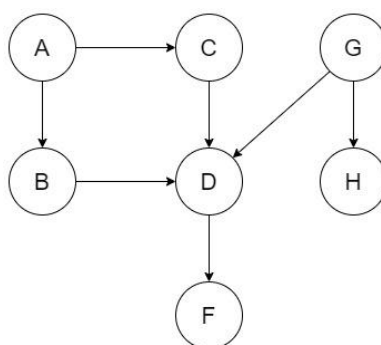
توجه: استفاده از موبایل، تبلت، لپ‌تاپ و هر گونه وسیله الکترونیکی غیر مجاز است.

سوال ۱: سوالات پاسخ کوتاه (۲۴ نمره)

درست یا نادرست بودن هر یک از گزاره‌های زیر را مشخص کرده و علت را به صورت مختصر توضیح دهید. (هر مورد ۴ نمره)

الف) با فرض صادق بودن faithfulness، با استفاده از الگوریتم PC می‌توان گراف جهت‌دار علّیت را به طور کامل بدست آورد.

ب) در مدل‌های VAE، اضافه کردن یک ترم به تابع هزینه که باعث شود توزیع احتمال فضای پنهان به سمت یک توزیع خاص (مثلاً نرمال استاندارد) میل کند، باعث معنادار شدن فضای پنهان می‌شود.



ج) در گراف بالا، $A \perp H$

د) در گراف بالا، $B \perp H | D, G$

ه) در گراف بالا، $p(A | do F = f) = p(A | F = f)$

و) در گراف بالا، $p(F | do C = c) = p(F | C = c)$

سوال ۲: شبکه‌های مارکوف و sampling (۱۸ نمره)

توزیع احتمال زیر را که روی یک گراف مارکوف تعریف شده است در نظر بگیرید:

$$p(x_1, \dots, x_5) = \frac{1}{Z} \phi_1(x_1, x_3, x_5) \phi_2(x_3, x_4) \phi_3(x_2, x_5) \phi_4(x_2, x_4)$$

الف) (۵ نمره) گراف بدون جهت مربوطه را با حداقل یال رسم کنید.

ب) (۳ نمره) اگر تابع پتانسیل ϕ_2 را ۵ برابر کنیم، توزیع احتمال حاصل چه تغییری می‌کند؟

ج) (۵ نمره) گراف جهت‌داری که Minimal I-Map گراف بالا باشد را بدست آورید.

د) (۵ نمره) چگونه می‌توانیم بدون داشتن ضریب نرمال کننده Z ، از توزیع بالا سمپل‌گیری کرد؟ توضیح دهید.

سوال ۳: مدل‌های علی (Counterfactual Question) (۱۰ نمره)

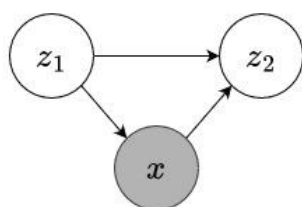
برای درمان بیماری کرونا، دو داروی A و B وجود دارد. افراد نسبت به این دو دارو، واکنش‌های مختلفی دارند و پزشکان هنوز علت تفاوت این واکنش‌ها را متوجه نشده‌اند. برخی افراد نسبت به هر دو دارو واکنش مثبت نشان می‌دهند و برخی نسبت به هر دو واکنش منفی دارند. برخی دیگر نسبت به یکی از این دو دارو واکنش منفی و نسبت به دیگری واکنش مثبت دارند. یک مطالعه آماری روی افراد مختلف انجام شده است که نتایج آن به شرح زیر است:

- ۱۰ درصد افراد نسبت به هر دو دارو واکنش منفی دارند.
- ۵۰ درصد افراد نسبت به هر دو دارو واکنش مثبت دارند.
- ۲۵ درصد افراد فقط نسبت به داروی A واکنش منفی دارند.
- ۱۵ درصد افراد فقط نسبت به داروی B واکنش منفی دارند.

فردی به پزشک مراجعه می‌کند و پزشک داروی B را برای او تجویز می‌کند. متأسفانه فرد بیمار فوت می‌کند 😞. پزشک می‌خواهد بداند که اگر داروی A را برای این فرد تجویز می‌کرد، با چه احتمالی این فرد درمان می‌شد؟

سوال ۴: Variational Inference (Mean-Field) (۲۰ نمره)

شبکه‌ی بیزی زیر را در نظر بگیرید که در آن متغیر x مشاهده شده و متغیرهای z_1 و z_2 پنهان هستند. با استفاده از روش Mean-Field تقریب توزیع احتمال پسین $p(z_1, z_2 | x)$ را بدست آورید.



$$p(z_1) = \lambda e^{-\lambda z_1}$$

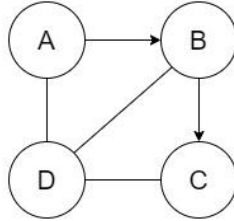
$$p(x|z_1) = N(x|z_1, 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x - z_1)^2\right)$$

$$p(z_2|x, z_1) = N(z_2|x + z_1, 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(z_2 - x - z_1)^2\right)$$

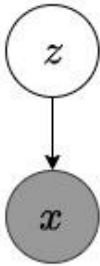
راهنمایی: هر دو توزیع $q(z_1)$ و $q(z_2)$ به صورت توزیع‌های معروف می‌شوند.

سوال ۵: مدل‌های علی (Causal Discovery) (۸ نمره)

احمد و سعید الگوریتم PC را روی چهار متغیر A, B, C, D اجرا کرده‌اند و گراف زیر بدست آمده است. احمد ادعا می‌کند که جهت یال CD را می‌توان تعیین کرد ولی سعید مخالف است. حق با کیست؟ چرا؟



سوال ۶: Variational Inference (۱۵ نمره)



شبکه بیزی روبرو را به همراه توزیع‌های شرطی زیر در نظر بگیرید.

$$p(z) = z e^{-\frac{z^2}{2}} \quad z \geq 0$$

$$p(x|z) = \mathcal{N}(x|z, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-z)^2}$$

می‌خواهیم با استفاده از Variational Inference، توزیع احتمال پسین $p(z|x)$ را تقریب بزنیم. برای این کار، توزیع پسین را با یک توزیع رایلی با پارامتر α به صورت زیر تقریب می‌زنیم:

$$q(z) = \frac{z}{\alpha} e^{-\frac{z^2}{2\alpha}}$$

مقدار بهینه پارامتر α را بدست آورید.

راهنمایی: برای یک متغیر تصادفی z با توزیع احتمال رایلی با پارامتر α داریم:

$$E[z] = \sqrt{\frac{\alpha\pi}{2}} \quad E[z^2] = 2\alpha$$

سوال ۷: مدل‌های مبتنی بر جریان (Normalizing Flow) (۱۵ نمره)

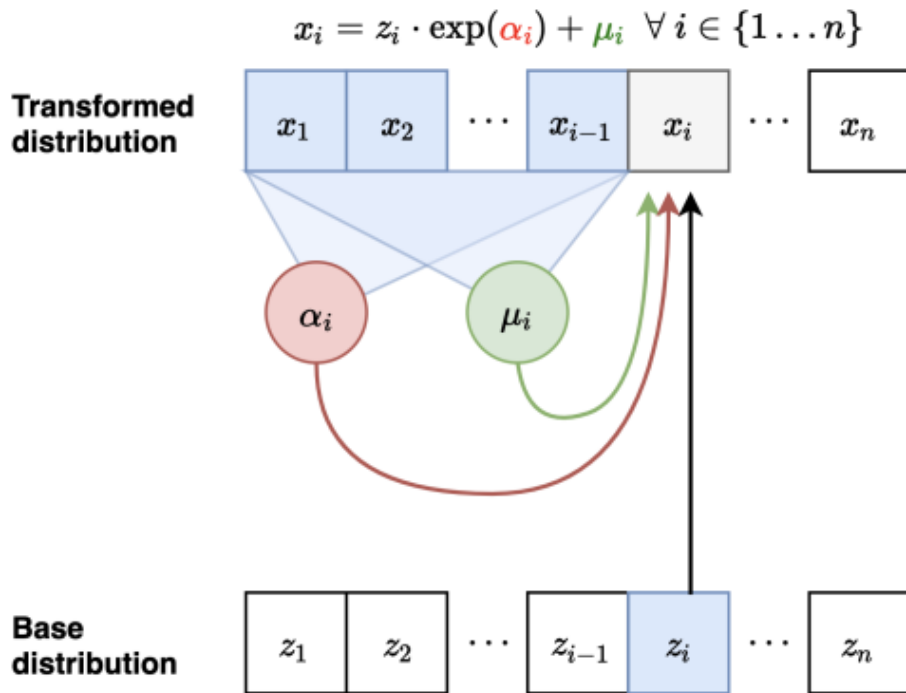
می‌خواهیم یک تبدیل از فضای $\mathbf{z} \in \mathbb{R}^n$ به فضای $\mathbf{x} \in \mathbb{R}^n$ به صورت زیر بدست آوریم:

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \xrightarrow{f} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

برای این منظور، از تبدیل زیر استفاده می‌کنیم:

$$x_i = z_i \exp(\alpha_i) + \mu_i \quad \forall i \in \{1, \dots, n\}$$

که در رابطه بالا، α_i و μ_i دو تا شبکه عصبی هستند که ورودی‌هایشان $[x_1, \dots, x_{i-1}]$ است و خروجی آن‌ها یک عدد اسکالر است (فرض کنید $x_1 = z_1$). همچنین فرض کنید، $\mathbf{z} \sim N(0, I)$ باشد. شکل زیر، به صورت شماتیک، تبدیل بالا را نشان می‌دهد.



الف) (۵ نمره) توضیح دهید که ساختار بالا، می‌تواند در یک شبکه‌ی normalizing flow استفاده شود.

ب) (۵ نمره) دترمینان ماتریس ژاکوبین تبدیل بالا را بدست آورید.

ج) (۵ نمره) مشکل ساختار بالا این است که برای تولید نمونه خیلی کند است. زیرا x_i ها را باید به ترتیب (به صورت sequential) تولید کرد. به نظر شما چگونه می‌توان این مشکل را حل کرد؟ به عبارت دیگر، چه تغییری در ساختار بالا می‌توان ایجاد کرد، تا سرعت تولید نمونه افزایش یابد.

موفق و پیروز باشید.