آمار و احتمال مهندسی

نیم‌سال اول ۱۴۰۳-۱۴۰۲

مدرس: دکتر نجفی

دانشکده‌ی مهندسی کامپیوتر

تمرین سری پنجم | موعد تحویل: ۱ دی

- مهلت ارسال پاسخ تا ساعت ۲۳:۵۹ روز مشخص‌شده است.
- هم‌کاری و هم‌فکری شما در انجام تمرین مانعی ندارد اما پاسخ ارسالی هر کس حتما باید توسط خود او نوشته شده باشد.
- در صورت هم‌فکری و یا استفاده از هر منابع خارج درسی، نام هم‌فکران و آدرس منابع مورد استفاده برای حل سوال مورد نظر را ذکر کنید.
- لطفا تصویری واضح از پاسخ سوالات نظری بارگذاری کنید. در غیر این صورت پاسخ شما تصحیح نخواهد شد.

مسئله‌ی ۱. (۱۵ نمره)

گروهی متشکل از ۲۰۰ نفر شامل ۱۰۰ مرد و ۱۰۰ زن داریم که به صورت تصادفی به ۱۰۰ تیم دونفره تقسیم شده‌اند. با استفاده از نامساوی چیبیشف، یک کران بالا برای احتمال این رویداد بیابید که حداکثر ۳۰ تیم هم عضو مرد و هم عضو زن داشته باشند.

حل. ابتدا به مردها به صورت دلخواه یک شماره از ۱ تا ۱۰۰ می‌دهیم و سپس متغیر تصادفی زیر را برای $i = 1, 2, \dots 100$ تعریف می‌کنیم:

$$X_i = \begin{cases} 1 & \text{اگر مرد } i \text{ با یک زن گروه شده است} \\ 0 & \text{موارد دیگر} \end{cases}$$

بنابراین X، تعداد گروه های مختلط به صورت زیر تعریف می شود:

$$X = \sum_{i=1}^{100} X_i$$

$$E[X_i] = P\{X_i = 1\} = \frac{100}{199}$$

به طور مشابه، برای هر $i \neq j$،

$$E[X_iX_j] = P\{X_i = 1, X_j = 1\} = P\{X_i = 1\}P\{X_j = 1 | X_i = 1\} = \frac{100}{199} \cdot \frac{99}{197}$$

پس $P\{X_j = 1 | X_i = 1\} = \frac{99}{197}$، بنابراین از آنجایی که مرد i با یک زن گروه شده است پس مرد j به همان اندازه محتمل است که با هر یک از ۱۹۷ انسان دیگر گروه شود که ۹۹ تای آنها زن هستند بنابراین داریم:

$$E[X] = \sum_{i=1}^{100} E[X_i] = (100) \cdot \frac{100}{199} \approx 50/25$$

$$\begin{aligned} Var(X) &= \sum_{i=1}^{100} Var(X_i) + 2 \sum_{i<j} Cov(X_i, X_j) \\ &= (100) \cdot \frac{100}{199} \cdot \frac{99}{199} + (2)\binom{100}{2} \left(\frac{100}{199} \cdot \frac{99}{197} - \left(\frac{100}{199}\right)^2 \right) \approx 25/126 \end{aligned}$$

طبق نامساوی Chebyshev

$$P\{X \leqslant ۳۰\} \leqslant P\{|X - ۵۰/۲۵| \geqslant ۲۰/۲۵\} \leqslant \frac{۲۵/۱۲۶}{(۲۰/۲۵)^۲} \approx /۰۶۱.$$

بنابراین احتمال آن کمتر از ۶ درصد است.

▷

مسئله‌ی ۲. (۱۱ نمره)

فرض کنید فرض کنید X و Y دو متغیر تصادفی مستقل $N(۰, ۱)$ باشند و

$$Z = ۱ + X + XY^۲$$

$$W = ۱ + X$$

مقدار $Cov(Z, W)$ را به دست آورید.

حل.

$$\begin{aligned}
Cov(Z, W) &= Cov(۱ + X + XY^۲, ۱ + X) \\
&= Cov(X + XY^۲, X) \\
&= Cov(X, X) + Cov(XY^۲, X) \\
&= Var(X) + E[X^۲Y^۲]E[XY^۲]E[X] \\
&= ۱ + E[X^۲]E[Y^۲]E[X]^۲E[Y^۲] \\
&= ۱ + ۱ - ۰ = ۲
\end{aligned}$$

▷

مسئله‌ی ۳. (۹ نمره)

فرض کنید $X_۱, X_۲, X_۳, ..., X_n$ نمونه‌ای تصادفی از توزیع هندسی (θ) است که در آن θ نامعلوم است. تخمینگر بیشینه احتمالی(MLE) برای θ را بر اساس این نمونه تصادفی به دست آورید.

حل.

$$X_i \sim Geometric(\theta)$$

$$P_{X_i}(x; \theta) = (۱ - \theta)^{x-۱}\theta.$$

$$\begin{aligned}
L(x_۱, x_۲, \cdots, x_n; \theta) &= P_{X_۱X_۲\cdots X_n}(x_۱, x_۲, \cdots, x_n; \theta) \\
&= P_{X_۱}(x_۱; \theta)P_{X_۲}(x_۲; \theta)\cdots P_{X_n}(x_n; \theta) = (۱ - \theta)^{\sum_{i=۱}^n x_i - n}\theta^n.
\end{aligned}$$

$$\ln L(x_۱, x_۲, \cdots, x_n; \theta) = (\sum_{i=۱}^n x_i - n)\ln(۱ - \theta) + n\ln\theta.$$

بنابراین

$$\frac{d\ln L(x_۱, x_۲, \cdots, x_n; \theta)}{d\theta} = (\sum_{i=۱}^n x_i - n) \cdot \frac{-۱}{۱ - \theta} + \frac{n}{\theta}.$$

بیشینه با ۰ قرار دادن مشتق به دست می آید پس

$$\hat{\theta}_{ML} = \frac{n}{\sum_{i=1}^{n} x_i}.$$

$$\hat{\Theta}_{ML} = \frac{n}{\sum_{i=1}^{n} X_i}.$$

$\triangleright$

مسئله‌ی ۴. (۸ نمره)

فرض کنید که X یک متغیر تصادفی و f, g دو تابع صعودی باشند. نشان دهید:

$$Cov(f(x), g(x)) \geqslant 0$$

حل.
فرض کنیم Y یک متغیر تصادفی هم توزیع و مستقل از X باشد. چون f, g صعودی هستند داریم:

$$(f(X) - f(Y))(g(X) - g(Y)) \geqslant 0$$

چون عبارت بالا همواره بزرگتر یا مساوی صفر است با امید گرفتن از آن نیز مقدار آن بزرگتر یا مساوی صفر خواهد بود:

$$\mathbb{E}(f(X) - f(Y))(g(X) - g(Y)) \geqslant 0$$

$$\Rightarrow \mathbb{E}(f(X)g(X)) - \mathbb{E}(f(Y)g(X)) - \mathbb{E}(f(X)g(Y)) + \mathbb{E}(f(Y)g(Y))$$

$$\Rightarrow \mathbb{E}(f(x)g(x)) - \mathbb{E}(f(Y)g(X)) \geqslant 0$$

چون X, Y مستقل هستند، داریم:

$$\mathbb{E}(f(Y)g(X)) = \mathbb{E}(f(Y))\mathbb{E}(g(X)) = \mathbb{E}(f(X))\mathbb{E}(g(X))$$

$$\Rightarrow \mathbb{E}(f(x)g(x)) - \mathbb{E}(f(X))\mathbb{E}(g(X)) \geqslant 0$$

$$\Rightarrow Cov(f(X), g(X)) \geqslant 0$$

$\triangleright$

مسئله‌ی ۵. (۱۶ نمره)

در محله طرشت یک دستگاه خودپرداز متعلق به بانک ملی وجود دارد که به صورت توزیع زیر پول نقد پرداخت می‌کند.

Amount, x (Toman)	50	100	200
$P(X = x)$	0.3	0.5	0.2

تعداد شهروندان طرشتی که روزانه به این دستگاه مراجعه می‌کنند از توزیع $N \sim Poisson(\lambda)$ پیروی می‌کند. فرض کنید $T_N = X_۱ + X_+ ... + X_N$ مقدار کل پول‌های نقدی باشد که شهروندان در یک روز از دستگاه دریافت می‌کنند در شرایطی که هر کدام از X_i ها از توزیعی که بالاتر تعریف کردیم پیروی می‌کنند و از یکدیگر مستقل هستند، علاوه بر این از N نیز مستقل هستند.

۱. نشان دهید $E(X) = ۱۰۵, Var(X) = ۲۷۲۵$

۲. $E(T_N)$ و $Var(T_N)$ را بیابید.

حل. ۱) بدیهی است کافیست طبق تعریف واریانس و امید ریاضی پیش بروید.

$$E(X) = ۰/۳ \times ۵۰ + ۰/۵ \times ۱۰۰ + ۰/۲ \times ۲۰۰$$

و برای واریانس داریم

$$Var(X) = ۰/۳ \times (۱۰۵ - ۵۰)^۲ + ۰/۵ \times (۱۰۵ - ۱۰۰)^۲ + ۰/۲ \times (۱۰۵ - ۲۰۰)^۲$$

۲) در این مساله داریم:

$$T_N = \Sigma_{i=1}^{N} X_i$$

اگر می‌دانستیم که تعداد ترم‌ها چندتا است به راحتی به $E(T_N)$ و $Var(X)$ به عنوان میانگین و واریانس جمع یک سری متغیر تصادفی مستقل از هم می‌رسیدیم. به همین منظور فرض می‌کنیم که می‌دانیم که تعداد ترم‌هایی که در حال جمع شدن هستند چندتا است. طبق استقلالی که در صورت سوال آمده است داریم.

$$E(T_N|N) = E(X_1 + ... + X_N|N) = E(X_1 + ... + X_N)$$

حال بدلیل استقلال متغیرها می‌توانیم بنویسیم:

$$= E(X_1) + ... + E(X_N)$$

مقدار N را به عنوان یک ثابت در نظر می‌گیریم.

$$N \times E(X) = 105N$$

بطور مشابه داریم.

$$\begin{gathered} Var(T_N|N) = 2725N \\ \to E(T_N) = E_N\{E(T_N|N)\} \\ = E_N(105N) = 105E_N(N) = 105\lambda \end{gathered}$$

در اینجا از خواص توزیع پواسون استفاده کردیم.

$$N \sim Poison(\lambda) \to E(N) = \lambda$$

بطور مشابه برای واریانس هم داریم.

$$
\begin{aligned}
Var(T_N) &= E_N\{Var(T_N|N)\} + Var_N\{E(T_N|N)\}\\
&= E_N\{2725N\} + Var_N\{105N\}\\
&= 2725E_N(N) + 105^2 Var_N(N)\\
&= 2725\lambda + 11025\lambda\\
&= 13750\lambda
\end{aligned}
$$

$\triangleright$

مسئله‌ی ۶. (۱۱ نمره)

فرض کنید $x_1, x_2, ..., x_n$ نمونه‌های متغیرهای تصادفی مستقل با توزیع یکسان(iid) از توزیعی به شکل زیر هستند.

$$
f_X(x|\theta) = \begin{cases} \frac{\theta x^{\theta-1}}{3^\theta}, & 0 \leqslant x \leqslant 3 \\ 0, & otherwise \end{cases}
$$

MLE پارامتر θ را بیابید.

حل.

$$
\begin{aligned}
L(x_1, ..., x_n|\theta) &= \prod_{i=1}^n \frac{\theta x_i^{\theta-1}}{3^\theta}\\
\ln L(x_1, ..., x_n|\theta) &= \Sigma_{i=1}^n (ln\theta + (\theta-1)\ln x_i - \theta ln3)\\
\frac{\partial}{\partial\theta} \ln L(x_1, ..., x_n|\theta) &= \Sigma_{i=1}^n (\frac{1}{\theta} + \ln x_i - \ln 3) = 0\\
\frac{n}{\theta} + \Sigma_{i=1}^n \ln x_i - n\ln 3 &= 0\\
\frac{n}{\hat{\theta}} &= n\ln 3 - \Sigma_{i=1}^n \ln x_i\\
\hat{\theta}_{MLE} &= \frac{n}{n\ln 3 - \Sigma_{i=1}^n \ln x_i}
\end{aligned}
$$

چک می‌کنیم که تخمینی که بدست آورده‌ایم ماکزیمم باشد با استفاده از چک کردن این که به ازای تمام مقادیر ممکن θ مشتق دوم منفی باشد.

$$
\frac{\partial^2}{\partial\theta^2} \ln L(x_1, ..., x_n|\theta) = \Sigma_{i=1}^n (-\frac{1}{\theta^2}) = -\frac{n}{\theta^2} < 0
$$

$\ln L(x_1, ..., x_n|\theta)$ is concave downward everywhere.

$\triangleright$

مسئله‌ی ۷. (۱۲ نمره)

فرض کنید یک مدل داریم که برای ورودی x_i خروجی y_i را با استفاده از رابطه‌ی زیر تولید می‌کند.

$$
y_i = ax_i + \epsilon_i
$$

به صورتی که ϵ_i در واقع نویز سیستم ما می‌باشد و از توزیع $\epsilon_i \sim N(0, \sigma^2)$ با پارامتر ثابت σ تبعیت می‌کنید. با فرض اینکه n داده به صورت $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$ داریم. با استفاده از MLE، a را تخمین بزنید.

حل. با توجه به اینکه ارور از تابع نورمال پیروی می‌کند داریم که

$$
\epsilon_i \sim \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\epsilon_i^2}{2\sigma^2}\right)
$$

Likelihood برای یک مشاهده (x_i, y_i) برابر است با:

$$
L(a, \sigma^2 \mid x_i, y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - ax_i)^2}{2\sigma^2}\right)
$$

با توجه به اینکه هر مشاهده مستقل از دیگری است، در نتیجه برای Likelihood داریم:

$$L(a,\sigma^{۲} \mid \mathbf{x},\mathbf{y}) = \prod_{i=۱}^{n} \frac{۱}{\sqrt{۲\pi\sigma^{۲}}} \exp\left(-\frac{(y_i - ax_i)^{۲}}{۲\sigma^{۲}}\right)$$

با گرفتن لگاریتم از تابع بالا به Log Likelihood می‌رسیم که برابر است

$$l(a,\sigma^{۲} \mid \mathbf{x},\mathbf{y}) = -\frac{n}{۲}\log\left(۲\pi\sigma^{۲}\right) - \frac{۱}{۲\sigma^{۲}}\sum_{i=۱}^{n}(y_i - ax_i)^{۲}$$

با گرفتن مشتق از تابع بالا مینیمم آن را بدست می‌اوریم تا $\hat{a}$ را بدست آوریم.

$$\frac{\partial}{\partial a}\left(l(a,\sigma^{۲} \mid \mathbf{x},\mathbf{y})\right) = \frac{۱}{\sigma^{۲}}\sum_{i=۱}^{n} x_i\,(y_i - ax_i) \Longrightarrow \hat{a} = \frac{\sum\limits_{i=۱}^{n} x_i y_i}{\sum\limits_{i=۱}^{n} x_i^{۲}}$$

$\triangleright$

مسئله‌ی ۸. (۱۸ نمره)

در این مسئله می‌خواهیم مقدار انتگرال یک تابع را در بازه $[a,b]$ بدست آوریم.

$$I(f) = \int_a^b f(x)\mathrm{d}x$$

برای اینکار در نظر می‌گیریم که یک تابع توزیع داریم که احتمال PDF آن را با g نشان می‌دهیم و فقط در بازه $[a,b]$ می‌توان از آن داده گرفت. حال با n داده‌گیری از این توزیع مقدار $I(f)$ را با تخمینگر زیر تخمین می‌زنیم.

$$\hat{I}_n(f) = \frac{۱}{n}\sum_{i=۱}^{n}\frac{f(X_i)}{g(X_i)}$$

۱. نشان دهید که $\mathbb{E}\left[\hat{I}_n\left(f\right)\right] = I\left(f\right)$

۲. نشان دهید که در تعداد داده بالا، مقدار تخمینگر ما به خود مقدار اصلی نزدیک می‌شود.

۳. مقدار $Var\left[\hat{I}_n\left(f\right)\right]$ را بدست آورید.

حل.

۱. با توجه به اینکه داریم هر داده از تابع توزیع احتمال g می‌آید در نتیجه داریم که:

$$\mathbb{E}\left[\frac{f(X)}{g(X)}\right] = \int_a^b \frac{f(x)}{g(x)}g(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x = I(f)$$
$$\Longrightarrow E\left[\hat{I}_n(f)\right] = \mathbb{E}\left[\frac{۱}{n}\sum_{i=۱}^{n}\frac{f(X_i)}{g(X_i)}\right] = \frac{۱}{n}\sum_{i=۱}^{n}\mathbb{E}\left[\frac{f(X_i)}{g(X_i)}\right] = I(f)$$

۲. با توجه به اینکه $\frac{f(X_i)}{g(X_i)}$ ها به هم وابسطه نیستند، طبق قانون اعداد بزرگ داریم که:

$$\hat{I}_n(f) = \frac{۱}{n}\sum_{i=۱}^{n}\frac{f(X_i)}{g(X_i)} \longrightarrow \mathbb{E}\left[\frac{f(X)}{g(X)}\right] = I(f)$$

۳. با توجه به اینکه $X \sim g$ در نتیجه $\sigma^۲$ را به صورت زیر تعریف می‌کنیم.

$$\sigma^۲ := \mathrm{Var}\left[\frac{f(X)}{g(X)}\right] = \mathbb{E}\left[\left(\frac{f(X)}{g(X)}\right)^۲\right] - \left(\mathbb{E}\left[\frac{f(X)}{g(X)}\right]\right)^۲ = \int_a^b \frac{f(x)^۲}{g(x)}\mathrm{d}x - (I(f))^۲$$

برای بدست آوردن واریانس نیز به این صورت عمل می‌کنیم:

$$\mathrm{Var}\left[\hat{I}_n(f)\right] = \mathrm{Var}\left[\frac{۱}{n}\sum_{i=۱}^{n}\frac{f(X_i)}{g(X_i)}\right] = \frac{۱}{n^۲}\sum_{i=۱}^{n}\mathrm{Var}\left[\frac{f(X_i)}{g(X_i)}\right] = \frac{۱}{n}\mathrm{Var}\left[\frac{f(X)}{g(X)}\right] = \frac{\sigma^۲}{n}$$

$\triangleright$

موفق باشید :)