به نام خدا

آنالیز احتمالاتی در ابعاد بالا (۱-۲۵۸۴۱)

تمرین سری اوّل

ترم پاییز ۰۵-۱۴۰۴

دانشکده‌ی مهندسی برق

دانشگاه صنعتی شریف

استاد: دکتر محمّدحسین یاسائی میبدی

مهلت تحویل: یک‌شنبه ۲۵ آبان ساعت ۲۳:۵۹

($*$) مسائلی که با ستاره مشخّص شده‌اند امتیازی هستند و حل‌کردن آن‌ها نمره‌ی امتیازی خواهد داشت!

۱ خواصّ متغیّرهای تصادفی زیرگوسی

در این تمرین تلاش می‌کنیم تعدادی گزاره که معادل زیرگوسی بودن هستند را اثبات کنیم.

۱. فرض کنید $\Phi : \mathbb{R}^{\geq 0} \mapsto \mathbb{R}$ تابعی صعودی و مشتق‌پذیر باشد. ثابت کنید:

$$\mathbb{E}\left[\Phi(|X|)\right] = \Phi(0) + \int_0^\infty \Phi'(t)\,\mathbb{P}\left[|X| \geq t\right]\,dt.$$

۲. ثابت کنید اگر X یک متغیّر تصادفی زیرگوسی با پارامتر σ و میانگین صفر باشد، آن‌گاه:

$$\mathbb{P}\left[|X| \geq t\right] \leq 2e^{\frac{-t^2}{2\sigma^2}}.$$

۳. ثابت کنید اگر X یک متغیّر تصادفی زیرگوسی با پارامتر σ و میانگین صفر باشد، آن‌گاه:

$$\mathbb{E}\left[e^{\frac{X^2}{6\sigma^2}}\right] \leq 2.$$

راهنمایی: از قسمت ۱ استفاده کنید.

۴. ثابت کنید اگر $\mathbb{E}\left[e^{\frac{X^2}{6\sigma^2}}\right] \leq 2$ آن‌گاه X یک متغیّر تصادفی زیرگوسی با پارامتر $\sigma\sqrt{18}$ است.

۵. ثابت کنید اگر X یک متغیّر تصادفی زیرگوسی با پارامتر σ و میانگین صفر باشد، آن‌گاه به ازای هر s که در شرط $|s| \leq \frac{1}{8\sigma^2}$ صدق کند، داریم:

$$\mathbb{E}\left[e^{s(X^2-\mathbb{E}[X^2])}\right] \leq \exp\left(\frac{32s^2\sigma^4}{1-4|s|\sigma^2}\right),$$

و نتیجه بگیرید که X^2 یک متغیّر تصادفی زیرنمایی است.

۲ خوشه‌ی بزرگ پنهان

در این مسئله قصد داریم یک الگوریتم تصادفی برای حلّ یک مسئله را توسعه داده و تحلیل کنیم.

فرض کنید n نفر در یک مهمانی شرکت کرده‌اند و هر دو نفر، مستقل از دیگران، با احتمال $\frac{۱}{۲}$ با یکدیگر دست می‌دهند. همچنین می‌دانیم که k نفر در این مهمانی حضور دارند که همگی یک‌دیگر را می‌شناسند، بنابراین این افراد با احتمال ۱ دو به دو با یکدیگر دست می‌دهند. هدف ما طراحی یک الگوریتم چندجمله‌ای است که بتواند تمام این k نفر را به درستی شناسایی کند. بدیهی است که هیچ الگوریتمی نمی‌تواند با اطمینان کامل این گروه را مشخص کند، از این‌رو به دنبال الگوریتمی هستیم که **با احتمال بالا** این کار را درست انجام دهد.

۱. برای شروع، مسئله را با گراف تصادفی $\mathscr{G}(n, \frac{۱}{۲}, k)$ مدل‌سازی می‌کنیم. در این مدل، بین هر دو رأس از n رأس موجود، به‌صورت مستقل از سایر یال‌ها، با احتمال $\frac{۱}{۲}$ یال قرار می‌دهیم. سپس مجموعه‌ای از k رأس را انتخاب می‌کنیم و در صورتی که بین آن‌ها یالی وجود نداشته باشد، آن را اضافه می‌کنیم تا این k رأس تشکیل یک خوشه کامل در گراف بدهند.

   فرض کنید یک نمونه گراف $X \sim \mathscr{G}(n, \frac{۱}{۲}, k)$ داده شده است. توزیع و امیدریاضی درجه‌ی هر رأس از گراف X را محاسبه کنید

۲. فرض کنید
$$k \geq c\sqrt{n \log n},$$
   برای یک ثابت $c > ۰$ به اندازه‌ی کافی بزرگ باشد. با استفاده از نتایج بخش قبلی، الگوریتمی طراحی کنید که بتواند با احتمال بالا تمام k رأس مربوط به گروه را شناسایی کند. نشان دهید که با افزایش $n \to \infty$، احتمال موفقیت این الگوریتم به ۱ میل می‌کند.

راهنمایی: از درجه‌ی رأس‌ها برای تمایز میان رأس‌های درون خوشه و بیرون از آن استفاده کنید. برای تحلیل احتمال خطا، از نامساوی‌های تمرکز نظیر نامساوی Hoeffding کمک بگیرید.

۳ تقویت الگوریتم‌های تصادفی

تصوّر کنید که الگوریتمی برای حلّ یک مسئله‌ی تصمیم‌گیری در اختیار داریم (به طور مثال این مسئله که عدد داده شده اول است یا خیر). فرض کنید که الگوریتم ما به طور تصادفی تصمیم می‌گیرد و با احتمال $\frac{۱}{۲} + \delta$ پاسخ صحیح می‌دهد، که تنها کمی بهتر از حدس زدن کاملاً تصادفی است. برای بهبود عملکرد الگوریتم، آن را N بار اجرا می‌کنیم و رأی اکثریت می‌گیریم. نشان دهید برای هر $\varepsilon \in (۰, ۱)$، پاسخ با احتمال حداقل $۱ - \varepsilon$ درست است، به شرطی که داشته باشیم:

$$N \geq \frac{۱}{۲\delta^۲} \ln\left(\frac{۱}{\varepsilon}\right).$$

۴ بیشینه‌ی متغیّرهای تصادفی زیرگوسی

فرض کنید $X_۱, X_۲, \ldots, X_n$ متغیّرهای تصادفی زیرگوسی با پارامتر مشترک σ باشند. توجّه کنید که این متغیّرهای تصادفی الزاما از هم مستقل نیستند. نشان دهید برای هر $\varepsilon > ۰$ داریم:

$$\mathbb{P}\left[\max_{i \leq n}\left\{X_i - \mathbb{E}[X_i]\right\} \geq (۱+\varepsilon)\sigma\sqrt{۲\log(n)}\right] \xrightarrow{n\to\infty} ۰.$$

راهنمایی: از کران اجتماع استفاده کنید:

$$\mathbb{P}\left[\max\{X, Y\} \geq t\right] = \mathbb{P}\left[X \geq t \vee Y \geq t\right] \leq \mathbb{P}[X \geq t] + \mathbb{P}[Y \geq t].$$

این مسئله نشان می‌دهد که ماکسیمم $\max_{i \leq n}\left\{X_i - \mathbb{E}[X_i]\right\}$ از متغیّرهای تصادفی زیربرگوسی با پارامتر σ در بیشترین حالت از مرتبه‌ی $\sigma\sqrt{۲\log(n)}$ است. این مثال ساده‌ترین نمونه از نقش اساسی کران‌های دنباله‌ای در تخمین اندازه‌ی ماکسیمم متغیّرهای تصادفی است.

۵ کران بالای چفت روی دم دوجمله‌ای

دنباله‌ی متغیّرهای تصادفی برنولی مستقل و هم‌توزیع $\{X_i\}_{i=1}^{n}$ با پارامتر $\alpha \in (0, 1/2]$ را در نظر بگیرید و متغیّر تصادفی دوجمله‌ای Z_n را به صورت زیر تعریف کنید:

$$Z_n = \sum_{i=1}^{n} X_i.$$

هدف این تمرین به‌دست‌آوردن کران بالای چفت (tight) برای $\mathbb{P}[Z_n \leq \delta n]$ برای هر $\delta \in (0, \alpha)$ است.

۱. نشان دهید:

$$\mathbb{P}[Z_n \leq \delta n] \leq e^{-nD(\delta\|\alpha)},$$

که در آن:

$$D(\delta\|\alpha) := \delta \log \frac{\delta}{\alpha} + (1-\delta) \log \frac{(1-\delta)}{(1-\alpha)},$$

دیورژانس کولبک–لایبلر بین توزیع‌های برنولی با پارامترهای δ و α است.

۲. نشان دهید کران قسمت ۱ اکیداً بهتر از کران هوفدینگ برای هر $\delta \in (0, \alpha)$ است.

در ادامه برای احتمال $\mathbb{P}[Z_n \leq \delta n]$ به ازای هر مقدار درنظرگرفته‌شده‌ی $\delta \in (0, \alpha)$ کران پایینی هم ارائه می‌دهیم. برای این منظور مقدار $m = \lfloor n\delta \rfloor$ را برابر بزرگترین عدد صحیح کوچکتر یا مساوی $n\delta$ در نظر می‌گیریم. همچنین تعریف می‌کنیم $\tilde{\delta} = \frac{m}{n}$.

۱. نشان دهید:

$$\frac{1}{n} \log \mathbb{P}[Z_n \leq \delta n] \geq \frac{1}{n} \log \binom{n}{m} + \tilde{\delta} \log \alpha + (1-\tilde{\delta}) \log(1-\alpha).$$

۲. نشان دهید:

$$\frac{1}{n} \log \binom{n}{m} \geq \phi(\tilde{\delta}) - \frac{\log(n+1)}{n},$$

که در آن تابع آنتروپی باینری برابر است با:

$$\phi(\tilde{\delta}) = -\tilde{\delta} \log \tilde{\delta} - (1-\tilde{\delta}) \log(1-\tilde{\delta}).$$

راهنمایی: زمانی که Y متغیّر تصادفی دوجمله‌ای با پارامتر $(n, \tilde{\delta})$ باشد، مقدار احتمال $\mathbb{P}[Y = l]$ به ازای $l = m = \tilde{\delta} n$ بیشینه می‌شود.

۳. نشان دهید:

$$\mathbb{P}[Z_n \leq \delta n] \geq \frac{1}{n+1} e^{-nD(\delta\|\alpha)}.$$

۶ مربّعات زیرگوسی‌ها

۱. ثابت کنید اگر X یک متغیّر تصادفی زیرگوسی با پارامتر σ و میانگین صفر باشد، آن‌گاه به ازای هر $q \in \mathbb{N}$ داریم:

$$\mathbb{E}\left[X^{2q}\right] \leq 2(2\sigma^2)^q q!.$$

۲. فرض کنید $X_1, X_2, \ldots, X_n$ متغیّرهای تصادفی زیرگوسی با پارامتر σ و میانگین صفر باشند. ثابت کنید:

$$\mathbb{E}\left[\max_i |X_i|\right] \leq C\sigma\sqrt{\log n},$$

که $C > 0$ ثابت است.

راهنمایی: می‌توانید از نامساوی $\max_i |X_i| \leq \left(\sum_{i=1}^{n} |X_i|^p\right)^{\frac{1}{p}}$ که برای هر $p \geq 1$ برقرار است استفاده کنید.

۷ نامساوی Khintchine

فرض کنید $X_1, \ldots, X_n$ متغیّرهای تصادفی زیرگوسی مستقل با میانگین صفر و پارامتر واحد باشند. تعریف می‌کنیم:

$$\mathbf{X} = (X_1, \ldots, X_n)^\top.$$

اثبات کنید که برای هر بردار $\mathbf{a} = (a_1, \ldots, a_n)^\top \in \mathbb{R}^n$ و هر $p \in [2, \infty)$ داریم:

$$\sqrt{\mathbf{a}^\top \mathbf{a}} \leq \left(\mathbb{E}\left[|\mathbf{a}^\top \mathbf{X}|^p\right]\right)^{\frac{1}{p}} \leq CK\sqrt{p}\sqrt{\mathbf{a}^\top \mathbf{a}}.$$

که در آن:

$$K = \max_i \|X_i\|_{\psi_2} = \max_i \left\{\inf\left\{t > 0 : \mathbb{E}\left[\exp\left(\frac{X_i^2}{t^2}\right)\right] < 2\right\}\right\}$$

و C یک ثابت مطلق است.

۸ $(*)$ نامساوی Khintchine برای $p = 1$

نشان دهید با همان شرایط سوال قبل، داریم:

$$c(K\sqrt{\mathbf{a}^\top \mathbf{a}} \leq \mathbb{E}\left[|\mathbf{a}^\top \mathbf{X}|\right] \leq \sqrt{\mathbf{a}^\top \mathbf{a}}.$$

در اینجا $K = \max_i \|X_i\|_{\psi_2}$ و $c(K) > 0$ مقداری است که ممکن است فقط به K بستگی داشته باشد.

۹ $(*)$ گراف تصادفی تنک دشوار!

۱. گراف تصادفی اردوش–رنی را تصور کنید که متوسط درجه‌ی هر رأس $d = o(\log n)$ باشد. نشان دهید با احتمال بالا (مثلاً ۰٫۹) راسی وجود دارد که درجه‌اش از مرتبه‌ی $10d$ باشد.

۲. گراف تصادفی اردوش–رنی را تصور کنید که متوسط درجه‌ی هر رأس $d = \mathcal{O}(1)$ باشد. نشان دهید با احتمال بالا (مثلاً ۰٫۹) راسی با درجه‌ای حداقل از مرتبه‌ی $\mathcal{O}\left(\frac{\log n}{\log\log n}\right)$ وجود دارد.