دانشکده‌ی مهندسی برق

آنالیز احتمالاتی در ابعاد بالا		زمستان ۱۴۰۱
	تمرین سری اول	
مدرس: دکتر محمد حسین یاسائی		موعد تحویل: ۵ فروردین

۱ مقایسه‌ی دو نوع آماره

فرض کنید $X \geq 0$ یک متغیر تصادفی بوده و تابع مولد گشتاور آن در یک بازه حول صفر موجود باشد. نشان دهید:

$$\inf_{k=0,1,2,\cdots} \frac{\mathbb{E}[|X|^k]}{\delta^k} \leq \inf_{\lambda>0} \frac{\mathbb{E}[e^{\lambda X}]}{e^{\lambda\delta}}.$$

از نامساوی بالا چه نتیجه‌ای می‌گیرید؟

۲ خواص متغیرهای زیرگوسی

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

(آ) فرض کنید Φ تابعی صعودی و مشتق‌پذیر باشد. ثابت کنید:

$$\mathbf{E}[\Phi(|X|)] = \Phi(0) + \int_0^\infty \Phi'(t)\mathbf{P}[|X| \geq t]dt$$

از این اتحاد در قسمت‌های بعدی استفاده خواهیم کرد.

در ادامه فرض کنید X متغیر تصادفی با میانگین صفر باشد. دقت کنید که ثابت‌های عددی که در این مسئله می‌بینید بهترین نیستند و چندان اهمیتی هم ندارد که بهترین باشند! صرفاً وجود آنها برای ما مهم است.

(ب) ثابت کنید اگر $\mathrm{P}[|X| \geq t] \leq 2e^{-t^2/2\sigma^2}$ آنگاه $\mathbf{E}\left[e^{X^2/6\sigma^2}\right] \leq 2$.

راهنمایی: از قسمت (آ) استفاده کنید.

(ج) ثابت کنید اگر $\mathbf{E}\left[e^{X^2/6\sigma^2}\right] \leq 2$ آنگاه X زیرگوسی با پارامتر $18\sigma^2$ است.

راهنمایی: برای مقدارهای بزرگ λ از نامساوی $|\lambda X| \leq \frac{a\lambda^2}{2} + \frac{X^2}{2a}$ با انتخاب مقدار مناسبی برای a استفاده کنید و برای مقادیر کوچک λ از همان نامساوی به همراه $\mathbf{E}\left[e^{\lambda X}\right] \leq 1 + \frac{\lambda^2}{2}\mathbf{E}\left[X^2 e^{|\lambda X|}\right]$ که با بسط تیلور به سادگی اثبات می‌شود استفاده نمایید.

(د) ثابت کنید اگر X زیرگوسی با پارامتر σ باشد، آنگاه $\mathbf{E}\left[X^{2q}\right] \leq 2\left(2\sigma^2\right)^q q!$ برای هر $q \in \mathbb{N}$.

راهنمایی: از قسمت (آ) استفاده کنید.

(ه) نشان دهید اگر X زیرگوسی باشد، آنگاه $\forall |s| \leq \frac{1}{2\sigma^2}$: $\mathbb{E}\left[e^{s\left(x^2-\mathbb{E}x^2\right)}\right] \leqslant \exp\left(\frac{8\sigma^4 s^2}{1-2\sigma^2 s}\right)$ و نتیجه بگیرید X^2 زیرنمایی است.

۳ میانگین و واریانس متغیر زیرگوسی

متغیر تصادفی X را که در رابطه‌ی زیر صدق می‌کند، درنظر بگیرید:

$$\mathbb{E}[e^{\lambda X}] \leq e^{\frac{\lambda^2\sigma^2}{2}+\lambda\mu} \quad \forall\lambda \in \mathbb{R}$$

(آ) نشان دهید $\mathbb{E}[X] = \mu$.

(ب) نشان دهید $\mathrm{Var}(X) \leq \sigma^2$.

(ج) فرض کنید کوچکترین σ که در نامساوی بالا صدق می‌کند را بیابیم. آیا $\mathrm{Var}(X) = \sigma^2$ برقرار است؟ ثابت کنید یا مثال نقض بزنید.
راهنمایی: قرار دهید $g(\lambda) = \log \mathbb{E}[e^{\lambda X}] - \frac{\lambda^2\sigma^2}{2} - \lambda\mu$ و مقدار ماکزیمم تابع g را یافته و شرایط رسیدن به آن را بررسی کنید.

۴ بِنت و برن‌اشتاین

در این تمرین، به اثبات نامساوی بنت و برن‌اشتاین می‌پردازیم.

الف. فرض کنید برای یک $b \geq 0$ داشته باشیم $|X_i - \mathbb{E}X_i| \leq b$. نشان دهید که برای هر λ حقیقی

$$\log \mathbb{E}[e^{\lambda(X_i - \mathbb{E}X_i)}] \leq \sigma_i^2\lambda^2\Big[\frac{e^{\lambda b} - 1 - \lambda b}{(\lambda b)^2}\Big].$$

که در آن $\sigma_i^2 = \mathrm{var}(X_i)$. برای این اثبات، سعی کنید از بسط تیلور تابع نمایی بهره ببرید.

ب. فرض کنید $X_1, \ldots, X_n$ متغیرهای تصادفی مستقل بوده و همگی شرط بخش قبل را ارضا کنند. تعریف کنید $\sigma^2 = \frac{1}{n}\sum_{i=1}^n \sigma_i^2$. اثبات کنید

$$\mathbb{P}\Big[\sum_{i=1}^n (X_i - \mathbb{E}X_i) \geq n\delta\Big] \leq \exp\Big(-\frac{n\sigma^2}{b^2}h\big(\frac{b\delta}{\sigma^2}\big)\Big).$$

که $h(t) = (1 + t\log(1+t)) - t$ است.

ج. نشان دهید که برای یک انتخاب دیگر تابع h، کران بخش قبل، کران برن‌اشتاین خواهد بود. به کمک این واقعیت، نشان دهید که کران بنت، از برن‌اشتاین قوی‌تر است.

۵ کران زیر نمایی برای فرم مربعی ماتریس

فرض کنید ماتریس $Q \in \mathbb{R}^{n\times n}$، که ماتریسی مثبت نیمه‌معین است به شما داده شده است. همچنین فرض کنید $\{X_i\}_{i=1}^n$ متغیرهای تصادفی i.i.d. با توزیع نرمال استاندارد باشند. متغیر تصادفی Z را در نظر بگیرید:

$$Z = \sum_{i=1}^n \sum_{j=1}^n Q_{ij} X_i X_j$$

نشان دهید که اعداد ثابت c_1 و c_2 وجود دارند، به نحوی که

$$\mathbb{P}[Z \geq \mathrm{trace}(Q) + t] \leq 2e^{-\min\left(\frac{c_1 t}{||Q||_2}, \frac{c_2 t^2}{||Q||_{\mathrm{Fr}}^2}\right)}$$

که در آن، $||.||_2$، نرم اپراتوری ماتریس و $||.||_{\mathrm{Fr}}$، نرم فروبنیوس ماتریس است.

راهنمایی: به تجزیه‌ی مقدار ویژه‌ای ماتریس Q توجه کنید و از کران‌های زیرنمایی بهره ببرید.

۶ کران پایین!

۱. نشان دهید اگر Y یک متغیر تصادفی نامنفی باشد، آنگاه برای هر $a \in (0,1)$ داریم:

$$\mathbb{P}\{Y \geq a\mathbb{E}Y\} \geq (1-a)^2 \frac{(\mathbb{E}Y)^2}{\mathbb{E}[Y^2]}.$$

۲. فرض کنید متغیر تصادفی‌های $X_1, \ldots, X_n$ مستقل و با احتمال برابر مقادیر $\{0,1\}$ را می‌گیرند. نشان دهید،

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i > \frac{\sqrt{n}}{2}\right] \geq \frac{3}{32}.$$

۷ بزرگترین زیردنباله مشترک

فرض کنید $\mathbf{a} = a_1 \ldots a_n$ و $\mathbf{b} = b_1 \ldots b_n$ دو دنباله دودویی به طول n باشند که ارقام $\mathbf{a}, \mathbf{b}$ به صورت مستقل و یکنواخت از $\{0,1\}$ انتخاب شده‌اند. متغیر تصادفی X_n را طول بزرگترین زیردنباله‌ی مشترک آنها بگیرید. ثابت کنید مقدار X_n حول میانگین آن متمرکز است، در واقع ثابت کنید:

$$P\Big(\,|X_n - \mathbb{E}\,[X_n]| \geq \delta\Big) \leq 2\exp\left(-\delta^2/8n\right)$$

یادداشت: محاسبه‌ی میانگین X_n یک مسئله‌ی باز است، اما می‌دانیم:

$$0.788071 \leq \lim_{n\to\infty} \mathbb{E}\frac{X_n}{n} = \gamma \leq 0.826280$$

۸ گراف تصادفی تنک

۱. گراف تصادفی اردوش-رنی را تصور کنید که متوسط درجه‌ی هر راس $d = \mathcal{O}(\log n)$ باشد. نشان دهید با احتمال بالا (مثلاً 0.9) درجه‌ی همه‌ی رئوس از مرتبه ی $\mathcal{O}(\log n)$ می‌باشد.

۲. اینبار گراف تصادفی اردوش-رنی را تصور کنید که متوسط درجه‌ی هر راس $d = \mathcal{O}(1)$ باشد. نشان دهید با احتمال بالا (مثلاً 0.9) درجه‌ی همه‌ی رئوس از مرتبه‌ی $\mathcal{O}(\frac{\log n}{\log\log n})$ می‌باشد.

۹ کران بالای چفت روی دُم دوجمله‌ای

دنباله‌ی متغیرهای برنولی مستقل و هم‌توزیع $\{X_i\}_{i=1}^n$ با پارامتر $\alpha \in (0, 1/2]$ را در نظر بگیرید. حال متغیر تصادفی دوجمله‌ای $Z_n = \sum_{i=1}^n X_i$ را در نظر بگیرید. هدف این تمرین به دست آوردن کران بالای چفت (sharp) برای هر $\delta \in (0,\alpha)$، روی احتمال دم $\mathbb{P}[Z_n \leq \delta n]$ است.

۱. نشان دهید $\mathbb{P}[Z_n \leq \delta n] \leq e^{-nD(\delta||\alpha)}$ به طوری که

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

دیورژانس کولبک-لیبلر بین توزیع‌های برنولی با پارامترهای δ و α است.

۲. نشان دهید کران قسمت الف، 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 درنظر گرفته شده‌ی $\delta \in (0,\alpha)$ 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،

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

به طوری که $\phi(\tilde{\delta}) = -\tilde{\delta}\ \log\tilde{\delta} - (1-\tilde{\delta})\ \log(1-\tilde{\delta})$ انتروپی باینری است.

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

۳. نشان دهید،

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

۱۰ بیضی‌گون!

۱. ابتدا یک بیضی‌گون d بعدی را تعریف کنید. یک بیضی‌گون با چه پارامترهایی مشخص می‌شود؟

۲. در کلاس گزاره‌هایی در مورد پوسته و استوای کره‌ی واحد d بعدی بیان شد. مشابهاً گزاره‌هایی که فکر می‌کنید در مورد بیضی‌گون درست است را نوشته و اثبات کنید.

۳. اگر از یک بیضی‌گون که مرکز ثقلش روی مبدا است، به طور یکنواخت و مستقل دو نقطه انتخاب کنیم و بردارهایی که از مبدا به این دو نقطه وصل می‌شوند را در نظر بگیریم، زاویه این دو بردار چقدر خواهد بود؟ آیا می‌توانید گزاره‌ی دقیقی در مورد آن بگویید؟ در مورد متوسط طول پاره‌خطی که دو نقطه تصادفی را به هم متصل می‌کند چه می‌توان گفت؟ اگر n بردار تصادفی به این شکل انتخاب کنیم، مشابه حرفی که در کلاس برای کره زدیم و زاویه‌های دو به دو را کنترل می‌کردیم را می‌توان در مورد بیضی‌گون زد؟

۱۱ تقویت الگوریتم‌های تصادفی

تصور کنید که الگوریتمی برای حل یک مسئله‌ی تصمیم‌گیری در اختیار داریم (به طور مثال این مسئله که عدد داده شده اول هست یا خیر)، فرض کنید که الگوریتم ما به طور تصادفی تصمیم می‌گیرد و با احتمال $\frac{1}{2} + \delta$ پاسخ صحیح می‌دهد، که تنها کمی بهتر از حدس زدن کاملاً تصادفی است. برای بهبود عملکرد الگوریتم، آن را N بار اجرا می‌کنیم و رای اکثریت می‌گیریم. نشان دهید برای هر $\epsilon \in (0,1)$ پاسخ با احتمال حداقل $1-\epsilon$ درست است، به شرطی که داشته باشیم $N \geq (1/2)\delta^{-2}\ln\left(\varepsilon^{-1}\right)$

۱۲ تخمین استوار میانگین

فرض کنید می‌خواهیم میانگین μ از یک متغیر تصادفی X را با کمک نمونه‌های $X_1, \cdots, X_N$ که به طور مستقل از توزیع X نمونه‌برداری شده‌اند، به دست آوریم. ما به دنبال تخمین εـدقیق هستیم. به این معنا که می‌خواهیم تخمین ما در بازه‌ی $(\mu-\varepsilon, \mu+\varepsilon)$ قرار گیرد.

۱. نشان دهید $N = O\left(\sigma^2/\varepsilon^2\right)$ برای محاسبه تخمین εـدقیق با احتمال حداقل $\frac{3}{4}$ کافی است. $(\sigma^2 = VarX)$

۲. نشان دهید $N = O\left(\log\left(\delta^{-1}\right)\sigma^2/\varepsilon^2\right)$ برای محاسبه تخمین ϵـدقیق با احتمال حداقل $1-\delta$ کافی است.

۱۳ (امتیازی) گراف تصادفی تنک دوشوار!

۱. گراف تصادفی اردوش-رنی را تصور کنید که متوسط درجه‌ی هر راس $d = o(\log n)$ باشد. نشان دهید با احتمال بالا (مثلاً 0.9) راسی وجود دارد که درجه‌اش از مرتبه‌ی $10d$ باشد.

۲. گراف تصادفی اردوش-رنی را تصور کنید که متوسط درجه‌ی هر راس $d = \mathcal{O}(1)$ باشد. نشان دهید با احتمال بالا (مثلاً 0.9) راسی با درجه‌ای حداقل از مرتبه‌ی $\mathcal{O}(\frac{\log n}{\log\log n})$ وجود دارد.