نمونه سوالات میان‌ترم اول
درس آمار و احتمال مهندسی

مدرس: امیر نجفی		دانشکده کامپیوتر، دانشگاه صنعتی شریف
آذر ۱۴۰۲

توضیحات:

* زمان کل امتحان اصلی ۱۰۵ دقیقه خواهد بود. همچنین، امتحان اصلی شامل ۵ سوال خواهد بود.
* در پایان همه سوالات، تعدادی راهنمایی و روابط معروف که «ممکن» است به آنها نیاز داشته باشید آورده شده‌اند. دقت کنید که برای حل برخی سوالات ممکن است به برخی و یا هیچکدام از راهنمایی‌ها نیازی پیدا نکنید.

سوال ۱ (۶ نمره):

در بندهای زیر، سناریوهای مجزایی از نحوه تعریف یک یا چند متغیر تصادفی را می‌بینید. سپس، گزاره‌هایی در مورد آنان گفته خواهد شد. مشخص کنید که کدام بندها گزاره‌های غلط و یا درست را بیان می‌کنند. در مورد هر بند نیز یک یا دو خط توضیحات دهید. دقت کنید که بندهای مختلف در مورد متغیرهای تصادفی مختلف صحبت می‌کنند و به یکدیگر ارتباطی ندارند.

الف) در صورتیکه متغیرهای تصادفی X, Y مستقل باشند، حتماً خواهیم داشت $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

ب) در صورتیکه داشته باشیم $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$، آنگاه متغیرهای تصادفی X, Y حتماً مستقل هستند.

ج) فرض کنید که متغیرهای تصادفی X, Y مشترکاً گاوسی باشند. می‌توانیم پارامترهای توزیع را طوری انتخاب کنیم تا داشته باشیم $\mathbb{E}\left(Y|X\right) = 2X^2 + 3$.

د) فرض کنید متغیرهای تصادفی X, Y هر کدام توزیع‌های حاشیه‌ای یکنواخت در بازه $\left[0,1\right]$ داشته باشند. در این صورت، توزیع توامان این دو متغیر نیز حتماً یک توزیع یکنواخت روی مربع واحد $\left[0,1\right]^2$ است.

ه) فرض کنید $X_1, X_2, \ldots$ یک دنباله i.i.d. از متغیرهای تصادفی با میانگین $\mu = 0$ و واریانس محدود باشد. در این صورت، واریانس توزیع متغیر تصادفی جدید $Z = \sum_{i=1}^{\infty} \frac{1}{i} X_i$ همگرا بوده و وجود خواهد داشت.

و) در بند ه) فرض کنید متغیرهای تصادفی $X_1, X_2, \ldots$ دیگر i.i.d. نبوده، ولی کماکان هم توزیع با میانگین $\mu = 0$ و واریانس محدود باشند. در این صورت، جواب شما به سوال بند ه) چه تغییری خواهد کرد؟

سوال ۲ (۴ نمره):

شما در یک مسابقه بخت‌آزمایی شرکت کرده‌اید. این مسابقه می‌تواند تا بی‌نهایت ادامه پیدا کند، و شما در هر مرحله (مستقل از سایر مراحل) با احتمال p یک میلیون تومان جایزه دریافت کرده، ولی با احتمال $1 - p$ خواهید باخت. در صورت باخت از مسابقه خارج شده و تمام جوایز کسب شده تا آن لحظه را نیز از دست می‌دهید. از طرف دیگر، پیش از وقوع یک باخت و در هر زمانی که اراده کنید می‌توانید جایزه جمع‌آوری شده تا آن لحظه را برداشته و از مسابقه خارج شوید.

بر خلاف اصرار مجری برنامه، واضح است که ادامه دادن بی‌حد این مسابقه کار به صلاحی نیست. چون بالاخره در مرحله‌ای خواهید باخت و چیزی عایدتان نمی‌شود. فرض کنید که تصمیم گرفته‌اید در صورت طی شدن m مرحله موفقیت‌آمیز جوایزتان را برداشته و خداحافظی کنید.

الف) در صورتیکه $m = 1$ را انتخاب کنید، میانگین جایزه‌ای که دریافت می‌کنید چند میلیون تومان خواهد شد؟
ب) فرض کنید $m = 2$ را انتخاب کرده باشید. حالا میانگین جایزه کسب شده توسط چقدر است؟
ج) واریانس جایزه کسب شده در حالت ب) را محاسبه کنید.
د) بهترین انتخاب m برای اینکه میانگین جایزه دریافتی شما بیشینه شود چقدر است؟

سوال ۳ (۳ نمره):

متغیر تصادفی X دارای توزیع نمایی با پارامتر $\lambda > 0$ را در نظر بگیرید. گشتاور چهارم X یا $m_4 = \mathbb{E}\left(X^4\right)$ را محاسبه کنید.

سوال ۴ (۴ نمره):

یک گراف n راسی را در نظر بگیرید. فرض کنید میان هر زوج رأس از میان $\binom{n}{2}$ زوج رأس ممکن از گراف، به صورت مستقل از سایر زوج‌ها، با احتمال p یک یال گذاشته و با احتمال $1 - p$ هیچ یالی نمی‌گذاریم. به چنین گراف‌های تصادفی، اصطلاحاً گراف Erdős–Rényi نیز اطلاق می‌گردد.

الف) راس شماره ۱ را در نظر بگیرید. به تعداد رئوسی که به این رأس خاص مستقیماً متصل هستند، درجه این رأس اطلاق می‌گردد. برای درجه رأس شماره ۱ در این گراف، توزیع احتمال چیست؟

ب) به زیرمجموعه‌ای از رئوس یک گراف به همراه یال‌های بالقوه میان آنها یک «زیرگراف» اطلاق می‌گردد. می‌دانیم که گراف مسئله ما یک زیرگراف $n - 1$ رأسی مشخص دارد که می‌دانیم تماماً متصل یا اصطلاحاً «هم‌بند» است. یعنی بین هر دو رأس آن، مسیری پیوسته وجود دارد که ممکن است از سایر رئوس نیز عبور کند. در این صورت، احتمال اینکه کل گراف هم‌بند باشد چقدر است؟

سوال ۵ (۴ نمره):

متغیر تصادفی X دارای توزیع یکنواخت در بازه $[0,B]$ است. مقدار B برای ما دانسته شده نیست و قصد داریم آن را تخمین بزنیم. برای اینکار n نمونه مستقل و تصادفی از X گرفته و بیشینه آنان را تحت عنوان تخمینگر $\hat{B}$ (که قرار است تخمینی از B باشد) در نظر می‌گیریم. بالطبع، هر چقدر n بزرگ‌تر باشد، تخمین بهتری خواهیم داشت. فرض کنید اعداد ϵ, δ مقادیری مثبت و کوچک باشند.

الف) احتمال اینکه در تخمین B به وسیله $\hat{B}$ دچار خطایی بیش از ϵ شویم، یعنی $|\hat{B} - B| > \epsilon$، چقدر است؟ (ممکن است در جواب شما مقدار نامعلوم B نیز ظاهر شود، که مشکلی ندارد.)

ب) یک کران پایین برای تعداد نمونه‌ها n را به صورت تابعی از ϵ, δ به گونه‌ای بیابید که بتوان تضمین کرد با احتمال حداقل $1 - \delta$، خطای تخمین، یعنی $|\hat{B} - B|$، از ϵ کوچک‌تر است.

سوال ۶ (۳ نمره):

متغیرهای تصادفی X, Y به صورت .i.i.d و دارای توزیع یکنواخت در بازه $[0,1]$ هستند. متغیرهای تصادفی جدید Z, W را به صورت زیر تعریف میکنیم:

$$Z = X + Y \quad , \quad W = X - Y$$

آیا Z و W از یکدیگر مستقل‌اند؟ در صورت استقلال، آن را اثبات کرده و در صورت عدم استقلال، حداقل با ذکر یک حالت خاص وابستگی را نشان دهید.

سوال ۷ (۴ نمره):

متغیرهای تصادفی $X_1, \ldots, X_n$ به صورت .i.i.d از چگالی احتمال $f_X(x)$ (و متناظراً توزیع انباشته احتمال $F_X(x)$) حاصل شده‌اند.
فرض کنید که از این متغیرهای تصادفی نمونه‌گیری شده و آنان را بر حسب مقدارشان به صورت صعودی مرتب کرده باشیم.
مقدار **یکی مانده به بزرگترین** در این دنباله به دلیلی برای ما مهم است. توزیع آماری آن را را بدست آورید.

توضیحات و راهنمایی‌ها:

* برای توزیع نمایی با پارامتر λ و $t < \lambda$ داریم: $M_X(t) = \mathbb{E}\left(e^{tX}\right) = \dfrac{\lambda}{\lambda - t}$

موفق باشید