به نام خدا

دانشگاه صنعتی اصفهان

دانشکده ریاضی

درس مبانی ریاضی علوم داده

ترم اول ۰۴-۰۵

تکلیف تئوری اول

۱. فرض کنید $X_1, X_2, \dots, X_n$ متغیرهای تصادفی مستقل و با توزیع یکسان باشند که میانگین آن‌ها μ و واریانس آن‌ها σ^2 است.

الف) اگر قرار دهیم $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. $E[\bar{X}]$ و $Var(\bar{X})$ را محاسبه کنید.

ب) فرض کنید n دانشجو در امتحان شرکت کرده‌اند و نمرات آن‌ها در بازه $[0,100]$ قرار دارد و میانگین نمرات برابر با ۵۰ است.

کران بالایی برای احتمال این‌که یک دانشجو نمره‌ای بیش از ۸۰ بگیرد به‌دست آورید.

پ) در ادامه بخش قبلی، فرض می‌کنیم که واریانس نمرات برابر با ۲۵ است. یک کران بالای دقیق‌تر برای احتمال این‌که یک دانشجو نمره‌ای بیش از ۸۰ بگیرد به‌دست آورید و پاسختان را با پاسخ بخش (ب) از نظر سرعت همگرایی و کاربرد مقایسه کنید.

ج) تعداد دانشجویان (n) باید چقدر باشد تا احتمال این‌که میانگین نمرات نمونه بیش از ۱۰ واحد از میانگین واقعی (۵۰) فاصله داشته باشد، حداکثر ۰/۰۱ باشد؟

۲. فرض کنید متغیر تصادفی X دارای توزیع زیر باشد :

$$p(x) = \begin{cases} 1 - \dfrac{1}{a}, & x = 0 \\ \dfrac{1}{a}, & x = a \end{cases}$$

الف) با استفاده از نابرابری مارکوف کران بالای احتمال $p(y \geq a)$ را برای $y = x, x^2, x^4$ محاسبه کنید.

ب) مقدار واقعی احتمال $p(y \geq a)$ را برای $y = x, x^2, x^4$ محاسبه کنید و با کران های به دست آمده در بخش قبلی مقایسه کنید همچنین تحلیل کنید که کران بالای مارکوف در چه شرایطی به مقدار اصلی احتمال نزدیکتر و در چه شرایطی دورتر می‌شود؟

۳. ابتدا بررسی کنید که حجم یک کره با شعاع ۲ چگونه با افزایش بعد فضا (d) تغییر می‌کند.
اگر شعاع کره مقداری بزرگتر از ۲ باشد، اما عددی ثابت و مستقل از بعد فضا، رفتار حجم چگونه خواهد بود؟
همچنین تعیین کنید شعاع r باید چه تابعی از d باشد تا حجم کره با افزایش بعد تقریبا ثابت باقی بماند.
راهنمایی: میتوانید برای ساده‌سازی محاسبات از تقریب استرلینگ برای فاکتوریل استفاده کنید ($n! \approx (\frac{n}{e})^n$).

۴. نمودار $V(d)$ را بر حسب d رسم کنید و مشخص کنید از چه مقدار d، حجم شروع به کاهش می‌کند؟

۵. ثابت کنید که به احتمال زیاد، زاویه‌ی بین دو بردار تصادفی در یک فضای با بعد بالا دست‌کم 45° است. راهنمایی: از قضیه ۲.۸ کتاب Hopcroft استفاده کنید، دقت کنید که بردارهای تصادفی در گوی یکه نیستند.

۶. خط استوای یک مکعب d بعدی را به صورت ابرصفحه ای تعریف کنید که در آن $\left\{x \middle| \sum_{i=1}^{d} x_i = \frac{d}{2}\right\}$ و مکعب واحد را مکعبی با اضلاعی به طول واحد و مرکز $(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2})$ تعریف می‌کنیم به نحوی که یکی ار رئوس آن در نقطه $(0,0,\dots,0)$ قرار دارد.

الف) آیا رئوس یک مکعب واحد نزدیک به خط استوا متمرکز شده اند؟

ب) آیا حجم مکعب واحد نزدیک به خط استوا متمرکز شده است؟

پ) آیا مساحت سطح مکعب واحد نزدیک به خط استوا متمرکز شده است؟

۷. اگر نقاطی در d بعد تولید شود که هر یک از مؤلفه‌ها یک توزیع گاوسی با واریانس واحد باشند، این نقاط به صورت تقریبی بر روی سطح یک کره به شعاع $\sqrt{d}$ قرار خواهند گرفت. توزیع آن‌ها، زمانی که نقاط بر روی یک خط تصادفی تصویر شوند را به دست آورید؟

۸. کره واحد A که مرکز آن در مبدأ مختصات قرار دارد و کره واحد B که مرکز آن در فاصله s از مبدأ قرار گرفته است را در نظر بگیرید. فرض کنید یک نقطه تصادفی x از توزیع ترکیبی زیر انتخاب شود:

"با احتمال $\frac{1}{2}$، نقطه را به صورت تصادفی از کره A انتخاب کنید و با احتمال $\frac{1}{2}$، نقطه را به صورت تصادفی از کره B انتخاب کنید."

نشان دهید که اگر فاصله مراکز دو کره (s) به اندازه کافی بزرگ باشد($s \gg \sqrt{d-1}$)، آنگاه این مقدار جدایی برای شرط $Prob(x \in A \cap B) = O(1)$ کافی است. به بیان دیگر، برای هر $\varepsilon > 0$ عدد ثابت c وجود دارد به طوری که اگر $s \geq c\sqrt{d-1}$ ، آنگاه

$.Prob(x \in A \cap B) < \varepsilon$

به عبارت ساده‌تر، این میزان جدایی به اندازه‌ای بزرگ است که تقریبا هیچ نقطه ای به طور همزمان در هر دو کره قرار نمی‌گیرد.

۹. فرض کنید X دارای توزیع گاوسی d بعدی با میانگین صفر و واریانس یک است ($X \sim N^d(0,1)$). میانگین و واریانس $\|X\|^2$ را بدست آورید.

راهنمایی: درباره گشتاور توزیع نرمال جست‌وجو کنید.

نکات تکمیلی :

- تکالیف فقط از طریق سامانه یکتا ارسال شود.
- در صورت مشاهده تقلب، نمره تکلیف صفر لحاظ می‌شود.
- در صورت وجود هرگونه ابهام سوالات خود را در گروه تلگرامی همین درس و یا از آیدی تلگرام @Setareh_Ghafouri بپرسید.