بسمه تعالی

دانشگاه صنعتی اصفهان

پاسخنامه تکلیف اول

مبانی ریاضی علوم داده

پاییز ۱۴۰۴

۱

- الف.

$$E[\bar{X}] = E[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} \times n\mu = \mu$$

برای محاسبه واریانس از فرض مستقل بودن متغیرها استفاده می‌کنیم.

$$Var(\bar{X}) = Var(\frac{1}{n}\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \frac{1}{n} X_i = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2} \times n\sigma^2 = \frac{\sigma^2}{n}$$

- ب. با استفاده از نامساوی مارکف این تقریب را انجام می‌دهیم.

$$P(X \geq a) \leq \frac{E[X]}{a} \to P(X \geq 80) = \frac{50}{80} = \frac{5}{8}$$

- پ. با استفاده از نامساوی چبیشف این تقریب را انجام می‌دهیم.

$$P(|X - E[X]| \geq a) = \frac{Var(x)}{a^2} \to P(|X - 50| \geq 30) = \frac{25}{30^2} = \frac{1}{36}$$

با مقایسه فرمول‌ها متوجه می‌شویم با فرض ثابت بودن واریانس مقدار تخمین زده شده در نامساوی چبیشف با سرعت a^2 کوچک می‌شود و نامساوی مارکف با سرعت a. همچنین نامساوی مارکف تنها با دانستن امید ریاضی X قابل استفاده است اما کران بدتری دارد. ولی نامساوی چبیشف درصورت داشت واریانس کران بهتری در اختیارمان می‌گذارد.

- ج. با استفاده از قضیه اعداد بزرگ این مسئله را حل می‌کنیم.

$$P(|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \geq \epsilon|) \leq \frac{\sigma^2}{n\epsilon^2}$$

$$P(|\bar{X} - 50| > 10) \leq \frac{25}{100n} \leq \frac{1}{100} \to n \geq 25$$

۲

- الف. برای $y = x^i$ مقدار تابع p به صورت

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

تعریف می‌شود. با توجه به مقادیر بالا مقدار $E[y] = 0 \times (1 - \frac{1}{a}) + a^i \times (\frac{1}{a}) = a^{i-1}$ می‌شود. حال برای جایگذاری در زنجیره مارکف کافی‌ست از رابطه $P(y \geq a) \leq \frac{E[y]}{a} = \frac{a^{i-1}}{a} = a^{i-2}$ استفاده کنیم. پس به ترتیب برای $y = x, x^2, x^4$ به ترتیب پاسخ برابر با $\frac{1}{a}, 1, a^2$ می‌باشد. لازم به ذکر است که برای $y = x^4$ می‌توان از کران 1 به جای a^2 استفاده کرد.

- ب. بدیهی است که مقدار واقعی این احتمال، در حالت تساوی رخ خواهد داد که تساوی با احتمال $\frac{1}{a}$ رخ می‌دهد. مقایسه آن با قسمت قبل این نتیجه را می‌دهد که برای $y = x^2, x^4$ مقدار تفاوت قابل توجهی در تخمین وجود دارد. اگر Y بتواند مقادیری بزرگتر از آستانه خود که در اینجا a هست بگیرد، ولی احتمال رخ دادن آن مقادیر بسیار کم یا ۰ باشد، آنگاه مقدار $E[y]$ افزایش پیدا می‌کند. درواقع در حالتی که بیشتر سهم $E[y]$ از مقادیر حول a بیاید، تخمین مارکف می‌تواند بهتر باشد.

۳

حجم کره dـبعدی واحد برابر با $V_d^1 = \frac{2 \cdot \pi^{d/2}}{d \cdot \Gamma(d/2)}$ می‌باشد. از طرفی حجم کره dـبعدی به شعاع r برابر با $r^d \times V_d^1$ می‌باشد. پس با گرفتن حد نتیجه را محاسبه می‌کنیم. (با فرض $r = 2$)

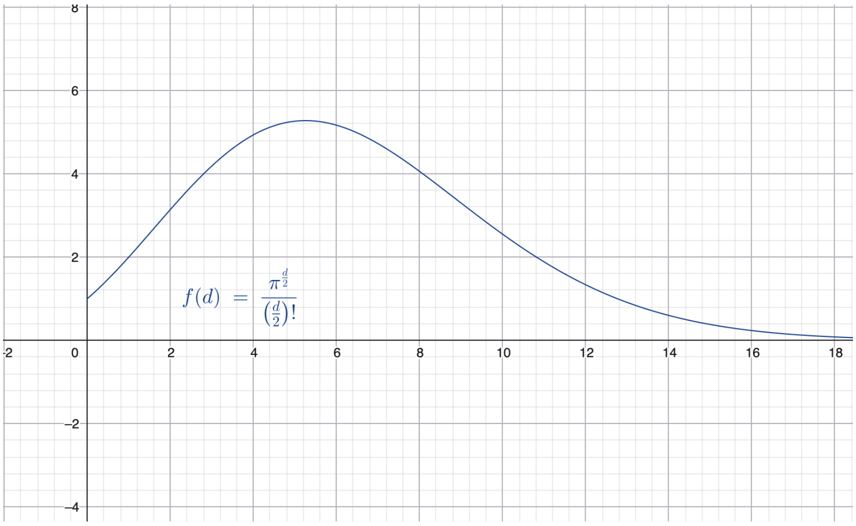
$$\lim_{d \to \infty} \frac{2 \cdot \pi^{d/2}}{d \cdot \Gamma(d/2)} r^d = \frac{(4 \cdot \pi)^{d/2}}{(d/2)!}$$

با محاسبه این مقدار خواهیم دید که تا $d/2 = 25$ مقدار تابع افزایش پیدا می‌کند، اما از آن پس شروع به کم شدن می‌کند و در $d/2 = 63$ به کمتر از ۱ می‌رسد.
درصورتی که مقدار $r \geq 2$ و ثابت باشد، باز هم حد تابع بالا به ۰ می‌رسد، اما d بزرگتر از حالت $r = d$ می‌شود. برای محاسبه مقدار r بر حسب d می‌توان از تقریب استرلینگ استفاده کرد.

$$n! = (\frac{n}{e})^n$$

$$\lim_{d \to \infty} \frac{r^d \pi^{d/2}}{(d/2)!} = \lim_{d \to \infty} \frac{r^d \cdot \pi^{d/2} \cdot (2e)^{d/2}}{d^{d/2}} \to r^d = \frac{d^{d/2}}{(2e\pi)^{d/2}} \to r = \sqrt{\frac{d}{2e\pi}}$$

۴



شکل ۱: نمودار حجم گوی یکه بر حسب بعد

برای $d \geq 6$ حجم شروع به کاهش می‌کند.

۵

استفاده از **قضیهٔ ۲٫۸ کتاب Foundations of Data Science** نشان بدهیم که در ابعاد بالا، زاویهٔ بین دو بردار تصادفی (نه لزوماً روی گوی یکه) با احتمال زیاد حداقل $45°$ است.

فرض کنید دو بردار تصادفی و مستقل

$$X, Y \in \mathbb{R}^d$$

داریم که توزیع آن‌ها **چرخش‌ناوردا** است؛ مثلاً:

- گاوسی استاندارد: $X, Y \sim \mathcal{N}(0, I_d)$، یا
- یکنواخت در گوی واحد (یا کرهٔ واحد).

زاویهٔ θ بین دو بردار با رابطه

$$\cos\theta = \frac{\langle X, Y\rangle}{\|X\|\,\|Y\|}$$

تعریف می‌شود. از آن‌جا که زاویه فقط به **جهت** بردارها بستگی دارد، طول‌ها مهم نیستند. بنابراین می‌توانیم بردارها را یکه کنیم:

$$U = \frac{X}{\|X\|}, \qquad V = \frac{Y}{\|Y\|}.$$

این دو بردار روی کرهٔ واحد S^{d-1} قرار می‌گیرند و چون توزیع اولیه چرخش‌ناورداست، جهت‌های U و V روی کره **یکنواخت** و مستقل‌اند. زاویهٔ بین X, Y همان زاویهٔ بین U, V است، چون

$$\cos\theta = \langle U, V\rangle.$$

پس مسئله را می‌توان به این شکل بازنویسی کرد:

> دو بردار مستقل U, V را به طور یکنواخت از کرهٔ واحد S^{d-1} انتخاب کنید. نشان دهید در ابعاد بالا با احتمال زیاد $\theta(U, V) \geq 45°$.

قضیهٔ ۲٫۸ کتاب بیان می‌کند که اگر n نقطه $x_1, \ldots, x_n$ را به طور مستقل و یکنواخت از **گوی واحد** در $\mathbb{R}^d$ بگیریم، آنگاه با احتمال حداقل $1 - O(1/n)$ دو خاصیت زیر هم‌زمان برقرارند:

۱. همهٔ نقاط تقریباً روی سطح گوی قرار دارند:

$$\|x_i\| \;\geq\; 1 - \frac{2\ln n}{d}, \qquad \forall i.$$

۲. هر دو نقطه تقریباً متعامدند:

$$|x_i \cdot x_j| \;\leq\; \sqrt{\frac{6\ln n}{d-1}}, \qquad \forall i \neq j.$$

حالا دو نقطهٔ متفاوت از میان این n نقطه انتخاب کنید؛ مثلاً x_1 و x_2. اگر رویداد «خوب» بالا رخ داده باشد، برای زاویهٔ θ بین این دو نقطه داریم:

$$\cos\theta = \frac{x_1 \cdot x_2}{\|x_1\|\,\|x_2\|}.$$

با استفاده از دو کران بالا می‌توانیم $|\cos\theta|$ را کران بالا بدهیم:

$$|x_1 \cdot x_2| \;\le\; \sqrt{\frac{6\ln n}{d-1}}, \qquad \|x_1\|, \|x_2\| \;\ge\; 1 - \frac{2\ln n}{d}.$$

پس:

$$|\cos\theta| \;=\; \frac{|x_1 \cdot x_2|}{\|x_1\|\,\|x_2\|} \;\le\; \frac{\sqrt{\dfrac{6\ln n}{d-1}}}{\left(1 - \dfrac{2\ln n}{d}\right)^2}.$$

انتخاب n و گذار به بعدهای بالا

حال n را به صورت یک تابع چندجمله‌ای از d انتخاب می‌کنیم؛ مثلاً:

$$n = d^2.$$

در این صورت:

$$\ln n = \ln(d^2) = 2\ln d.$$

جملهٔ صورت:

$$\sqrt{\frac{6\ln n}{d-1}} = \sqrt{\frac{12\ln d}{d-1}} = O\left(\sqrt{\frac{\ln d}{d}}\right),$$

که وقتی $d \to \infty$ می‌رود، به صفر میل می‌کند.

در مخرج داریم:

$$\left(1 - \frac{2\ln n}{d}\right)^2 = \left(1 - \frac{4\ln d}{d}\right)^2 \longrightarrow 1 \qquad (d \to \infty).$$

بنابراین کران کلی

$$|\cos\theta| \;\le\; \frac{\sqrt{\dfrac{12\ln d}{d-1}}}{\left(1 - \dfrac{4\ln d}{d}\right)^2}$$

وقتی d بزرگ می‌شود، به صفر میل می‌کند.

از این‌جا نتیجه می‌گیریم:

برای هر ثابت $c \in (0,1)$ (به‌خصوص $c = \frac{1}{\sqrt{2}} = \cos 45^\circ$)، عددی D وجود دارد که اگر $d \geq D$، آنگاه در رویداد خوب قضیهٔ ۲٫۸:

$$|\cos\theta| \;\leq\; c.$$

با انتخاب $c = \frac{1}{\sqrt{2}}$:

$$|\cos\theta| \;\leq\; \frac{1}{\sqrt{2}} \quad\Longrightarrow\quad \theta \in [45^\circ, 135^\circ].$$

پس به‌ویژه:

$$\theta \;\geq\; 45^\circ.$$

از آن‌جا که احتمال رویداد خوب حداقل $1 - O(1/n) = 1 - O(1/d^2)$ است، داریم:

$$\mathbb{P}\big(\theta \geq 45^\circ\big) \;\geq\; 1 - O\!\left(\frac{1}{d^2}\right),$$

که وقتی $d \to \infty$ می‌رود به ۱ میل می‌کند.

در ابعاد بالا، زاویهٔ بین دو نقطهٔ تصادفی در گوی (یا کره) واحد، با احتمال خیلی زیاد حداقل 45° است و در واقع به $90^\circ \approx$ نزدیک می‌شود.

«نه لزوماً در گوی یکه»

در صورت سؤال آمده بود که بردارها لزوماً در گوی یکه قرار ندارند. با این حال زاویه فقط به جهت بردار بستگی دارد، نه به اندازهٔ آن. اگر دو بردار تصادفی مستقل

$$X, Y \in \mathbb{R}^d$$

داشته باشیم که توزیع آن‌ها چرخش‌ناورداست (مثلاً گاوسی استاندارد)، جهت‌های نرمال‌شدهٔ آن‌ها

$$U = \frac{X}{\|X\|}, \qquad V = \frac{Y}{\|Y\|}$$

روی کرهٔ واحد S^{d-1} توزیعی یکنواخت دارند. زاویه بین X, Y برابر زاویه بین U, V است. بنابراین تمام استدلال بالا مستقیماً برای این حالت هم صدق می‌کند و نیازی نیست فرض کنیم که خود بردارها دقیقاً روی گوی واحد قرار دارند؛ کافی است توزیعشان چرخش‌ناوردا باشد.

به این ترتیب، بر اساس قضیهٔ ۲٫۸ کتاب، دیدیم که در فضای با بعد زیاد، «تقریباً همهٔ بردارهای تصادفی مستقل» تقریباً orthogonal هستند و زاویهٔ بین آن‌ها با احتمال زیاد حداقل 45° است.

۶

خط استوای مکعب dـبعدی را ابرصفحهٔ

$$H = \left\{ x \in \mathbb{R}^d \;\middle|\; \sum_{i=1}^{d} x_i = \frac{d}{2} \right\}$$

در نظر می‌گیریم. فاصلهٔ اقلیدسی یک نقطه x از این ابرصفحه برابر است با

$$\operatorname{dist}(x, H) = \frac{\left| \sum_{i=1}^{d} x_i - \frac{d}{2} \right|}{\sqrt{d}}.$$

(الف) آیا رؤوس مکعب واحد به خط استوا متمرکز می‌شوند؟

رؤوس مکعب مجموعهٔ $\{0,1\}^d$ هستند. اگر X یک رأس تصادفی باشد و

$$S = \sum_{i=1}^{d} X_i,$$

آنگاه هر X_i برنولی با پارامتر $1/2$ بوده و بنابراین

$$S \sim \operatorname{Bin}\big(d, \tfrac{1}{2}\big), \qquad \mathbb{E}[S] = \frac{d}{2}, \qquad \operatorname{Var}(S) = \frac{d}{4}.$$

فاصلهٔ رأس از استوا:

$$\operatorname{dist}(X, H) = \frac{|S - d/2|}{\sqrt{d}}.$$

قضیهٔ حد مرکزی می‌گوید:

$$\frac{S - d/2}{\sqrt{d/4}} \xrightarrow{d\to\infty} \mathcal{N}(0,1).$$

یعنی $|S - d/2| = \Theta(\sqrt{d})$ و بنابراین

$$\operatorname{dist}(X, H) = \Theta(1).$$

پس احتمال اینکه فاصله کمتر از ε باشد به عددی ثابت (وابسته به ε) میل می‌کند و برابر با ۱ نمی‌شود.

$\boxed{\text{رؤوس مکعب واحد به خط استوا متمرکز نمی‌شوند.}}$

(ب) آیا حجم مکعب واحد (نقاط داخل آن) به خط استوا متمرکز می‌شود؟

اگر

$$X = (X_1, \dots, X_d)$$

نقطه‌ای یکنواخت در $[0,1]^d$ باشد، آنگاه X_iها مستقل و

$$\mathbb{E}[X_i] = \frac{1}{2}, \qquad \mathrm{Var}(X_i) = \frac{1}{12}.$$

بنابراین

$$S = \sum_{i=1}^{d} X_i, \quad \mathbb{E}[S] = \frac{d}{2}, \qquad \mathrm{Var}(S) = \frac{d}{12}.$$

فاصله

$$\mathrm{dist}(X, H) = \frac{|S - d/2|}{\sqrt{d}}.$$

و طبق CLT:

$$\frac{S - d/2}{\sqrt{d/12}} \Rightarrow \mathcal{N}(0,1).$$

پس $|S - d/2| = \Theta(\sqrt{d})$ و بنابراین فاصله $\Theta(1)$ باقی می‌ماند و به صفر میل نمی‌کند.

$$\boxed{\text{حجم مکعب واحد نیز به خط استوا متمرکز نمی‌شود.}}$$

(پ) آیا مساحت سطح مکعب واحد به خط استوا متمرکز می‌شود؟

یک نقطهٔ تصادفی روی سطح مکعب (به جز لبه‌ها) روی یکی از $2d$ وجه قرار می‌گیرد. روی هر وجه، دقیقاً یک مؤلفه برابر 0 یا 1 و بقیه یکنواخت در $[0,1]$ هستند.

مدل نقطهٔ سطح:

$$S = B + \sum_{i=1}^{d-1} U_i,$$

که در آن $B \sim \mathrm{Bernoulli}(1/2)$ و $U_i \sim \mathrm{Unif}[0,1]$ مستقل‌اند.

پس

$$\mathbb{E}[S] = \frac{1}{2} + (d-1)\frac{1}{2} = \frac{d}{2},$$

$$\mathrm{Var}(S) = \frac{1}{4} + (d-1)\frac{1}{12} = \frac{d+2}{12}.$$

دوباره

$$\mathrm{dist}(X, H) = \frac{|S - d/2|}{\sqrt{d}},$$

و با CLT داریم

$$\frac{S - d/2}{\sqrt{(d+2)/12}} \Rightarrow \mathcal{N}(0,1).$$

پس فاصلهٔ نقاط روی سطح نیز $\Theta(1)$ بوده و به صفر میل نمی‌کند.

$\boxed{\text{مساحت سطح مکعب واحد هم به خط استوا متمرکز نمی‌شود.}}$

۷

فرض کنیم که یک نقطه X با بعد d داریم که توزیع گاوسی دارد. یعنی تمام مولفه‌های آن توزیع گاوسی دارند. حال یک خط مانند $< u_1, u_2, \dots, u_d >$ در نظر می‌گیریم که می‌خواهیم تصویر نقطه روی آن را محاسبه کنیم. با توجه به فرمول تصویر می‌دانیم $Proj_u(X) = (u^T X)u$ می‌باشد. که مقدار اسکالر آن برابر با $(u^T X)$ می‌باشد. پس برای نقطه X خواهیم داشت :

$$u^T \cdot X = x_1.u_1 + x_2.u_2 + \cdots + x_d.u_d$$

حال می‌دانیم که ترکیب خطی از متغیرهای گاوسی همچنان دارای توزیع نرمال می‌باشد (اثبات در فایل پیوست پاسخنامه قرار دارد).
حال برای پیدا کردن میانگین و واریانس این توزیع محاسبات زیر برقرار است.

$$E[u^T.X] = E[\sum_{i=1}^{d} x_i.u_i] = \sum_{i=1}^{d} E[x_i.u_i] = \sum_{i=1}^{d} E[x_i].u_i \xrightarrow{E[x_i]=0} E[u^T.X] = 0$$

و برای واریانس داریم

$$Var(u^T \cdot X) = u^T \cdot Var(X) \cdot u = u^T \cdot I_d \cdot u = ||u||^2$$

که با فرض یکه بودن بردار u برای اینکه فقط جهت مهم است می‌توان به واریانس ۱ نیز رسید.

۸

دو کرهٔ dـبعدی A و B با شعاع یکسان در نظر بگیرید که مرکز A در مبدأ و مرکز B در نقطهٔ

$$\mu = (s, 0, \dots, 0)$$

قرار دارد. نقطهٔ تصادفی X با توزیع ترکیبی زیر انتخاب می‌شود:

$$X = \begin{cases} \text{از یکنواخت نقطهٔ } A & \text{احتمال با } \tfrac{1}{2}, \\ \text{از یکنواخت نقطهٔ } B & \text{احتمال با } \tfrac{1}{2}. \end{cases}$$

به خاطر تقارن، داریم

$$\mathbb{P}(X \in A \cap B) = \tfrac{1}{2}\,\mathbb{P}(X \in A \cap B \mid X \in A) + \tfrac{1}{2}\,\mathbb{P}(X \in A \cap B \mid X \in B) = \mathbb{P}(X \in B \mid X \in A),$$

یعنی کافی است احتمال این‌که یک نقطهٔ یکنواخت از A در B هم قرار بگیرد را کران بزنیم.

مرحلهٔ ۱: تقریب «نقطهٔ یکنواخت در کره» با «نقطهٔ یکنواخت روی گوی»

از تمرین‌های قبل (تمرکز حجم در پوستهٔ نزدیک سطح) می‌دانیم که اگر d بزرگ باشد، تقریباً تمام حجم کره در پوسته‌ای بسیار نازک در نزدیکی سطح قرار دارد. بنابراین با خطای ناچیز (که بعداً در ε جذب می‌کنیم)، می‌توان فرض کرد X روی گوی به شعاع ثابت R (برای مثال $R = \sqrt{d-1}$) و به صورت یکنواخت توزیع شده است. از این به بعد X را یکنواخت روی گوی در نظر می‌گیریم.

مرحلهٔ ۲: شرط قرار گرفتن نقطه روی گوی در تقاطع دو کره

مرکز A در 0 و مرکز B در $\mu = (s, 0, \dots, 0)$ است. اگر $X = (X_1, \dots, X_d)$ روی گوی با شعاع R باشد، آنگاه

$$\|X\| = R.$$

شرط تعلق X به کرهٔ B این است که

$$\|X - \mu\| \le R.$$

با مجذور گرفتن و استفاده از $\|X\|^2 = R^2$ داریم:

$$\|X - \mu\|^2 = \|X\|^2 + \|\mu\|^2 - 2\langle X, \mu\rangle = R^2 + s^2 - 2sX_1 \le R^2.$$

پس

$$s^2 - 2sX_1 \le 0 \quad \Longrightarrow \quad X_1 \ \ge\ \frac{s}{2}.$$

بنابراین

$$\{X \in A \cap B\} \subseteq \{X_1 \geq s/2\}.$$

در نتیجه

$$\mathbb{P}(X \in A \cap B) \leq \mathbb{P}(X_1 \geq s/2).$$

مرحلهٔ ۳: توزیع X_1 روی گوی و تقریب گاوسی

برای بردار یکنواخت روی گوی dـ بعدی به شعاع $R = \sqrt{d-1}$، به خاطر چرخش‌ناوردایی، همهٔ مختصات توزیع یکسان دارند و

$$\mathbb{E}[X_1] = 0, \qquad \mathrm{Var}(X_1) = \frac{R^2}{d} = \frac{d-1}{d} \approx 1.$$

از طرفی می‌توان X را به صورت نرمال نرمال‌شده در نظر گرفت: اگر

$$Z \sim \mathcal{N}(0, I_d)$$

باشد، آنگاه

$$X = \frac{\sqrt{d-1}}{\|Z\|} Z$$

روی گوی به‌طور یکنواخت توزیع می‌شود و برای d بزرگ، ضریب $\frac{\sqrt{d-1}}{\|Z\|}$ تقریباً ۱ است. در نتیجه حاشیهٔ X_1 تقریباً گاوسی استاندارد است:

$$X_1 \approx Z_1, \qquad Z_1 \sim \mathcal{N}(0,1).$$

پس برای ثابت‌های مناسب و d به اندازهٔ کافی بزرگ، می‌توان نوشت

$$\mathbb{P}(X_1 \geq s/2) \approx \mathbb{P}(Z_1 \geq s/2),$$

و از کران دمی نرمال استفاده کرد:

$$\mathbb{P}(Z_1 \geq t) \leq e^{-t^2/2}, \qquad t > 0.$$

اگر

$$s \geq c\sqrt{d-1},$$

آنگاه

$$\mathbb{P}(X_1 \geq s/2) \;\lesssim\; \exp\left(-\frac{1}{2}\left(\tfrac{s}{2}\right)^2\right) \;\leq\; \exp\left(-\frac{c^2(d-1)}{8}\right).$$

این کران با بعد بزرگ به صورت نمایی کوچک می‌شود.

مرحلهٔ ۴: انتخاب c بر حسب ε

حال یک $\varepsilon > 0$ دلخواه در نظر بگیرید. کافی است c را آن‌قدر بزرگ انتخاب کنیم که

$$\exp\left(-\frac{c^2}{8}\right) \le \frac{\varepsilon}{2}.$$

مثلاً

$$c \ \ge \ \sqrt{8\ln\frac{2}{\varepsilon}}.$$

برای چنین cای و هر $d \ge 2$:

$$\exp\left(-\frac{c^2(d-1)}{8}\right) \le \exp\left(-\frac{c^2}{8}\right) \le \frac{\varepsilon}{2}.$$

از طرف دیگر، خطای ناشی از تقریب «کره به گوی» (یعنی نقاطی که شعاع آن‌ها به اندازهٔ قابل‌توجهی کمتر از R است) نیز طبق تمرکز شعاعی با احتمال نمایی کوچک رخ می‌دهد و می‌توان آن را هم حداکثر $\varepsilon/2$ در نظر گرفت (برای d به اندازهٔ کافی بزرگ).

در نتیجه برای $s \ge c\sqrt{d-1}$:

$$\mathbb{P}(X \in A \cap B) \ \le \ \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \ < \ \varepsilon.$$

$\boxed{\forall\, \varepsilon > 0\ \exists\, c > 0}$ به‌طوری‌که اگر $s \ge c\sqrt{d-1}$ باشد، آنگاه $\mathbb{P}(X \in A \cap B) < \varepsilon.$

به بیان ساده، اگر فاصلهٔ مراکز دو کره در مقیاس $\sqrt{d-1}$ به اندازهٔ کافی بزرگ باشد، آنگاه احتمال این‌که یک نقطهٔ تصادفی به طور هم‌زمان در هر دو کره قرار بگیرد، در ابعاد بالا تقریباً صفر است.

۹

ابتدا به محاسبه میانگین می‌پردازیم.
می‌دانیم که

$$E[||X^2||] = E[\sum_{i=1}^{d} x_i^2] = \sum_{i=1}^{d} E[x_i^2]$$

برقرار است. بعد از استناد به راهنمایی سوال درباره گشتاورها یا momentهای توزیع نرمال به رابطه زیر می‌رسیم.

$$E[X^p] = \begin{cases} 0 & \text{odd,} \\ \sigma^p \, (p-1)!! & \text{even} \end{cases}$$

$$(2k-1)! = \Pi_{i=1}^{(k)} (2i-1)$$

در نتیجه $E[x_i^2] = \sigma \times 1! = 1$ می‌شود.
پس :

$$E[||X^2||] = \sum_{i=1}^{d} 1 = d$$

برای محاسبه واریانس می‌توانیم از مستقل بودن متغیرهای بردار استفاده کنیم.

$$Var(||X^2||) = Var(||\sum_{i=1}^{d} x_i^2||) = \sum_{i=1}^{d} Var(x_i^2)$$

از طرفی می دانیم که

$$Var(x^2) = Var(x^4) - (Var(x^2))^2 = (3 \times 1) - 1 \rightarrow Var(x) = 2$$

پس داریم

$$\sum_{i=1}^{d} Var(x_i^2) = d \times 2 = 2d$$