IE 531

HW # 2

TIANQI WU

1. (a) $X \in \{-1, 0, 1\}$

$$Prob(X=-1) = Prob(X=1) = \frac{1}{2k^2}$$

$$X = \begin{cases} -1, & \text{Probability} = \frac{1}{2k^2} \\ 0, & \text{Probability} = 1 - \frac{1}{k^2} \\ 1, & \text{Probability} = \frac{1}{2k^2} \end{cases}$$

$$E(X) = -1 \times \frac{1}{2k^2} + 1 \times \frac{1}{2k^2} = 0$$

$$Var(X) = (-1)^2 \times \frac{1}{2k^2} + 1^2 \times \frac{1}{2k^2} = \frac{1}{k^2}$$

$$P(|X - E(X)| \geq \delta) \leq \frac{Var(X)}{\delta^2}$$

$$P(|X| \geq 1) = 2 \times \frac{1}{2k^2} = \frac{1}{k^2} \quad \text{when } \delta = 1$$

(b) Assume $E(X) = 0$. $Var(X) = 1$

Chebyshev tight bound: $P(|X| \geq \delta) = \frac{1}{\delta^2}$

when $X > 0$. CDF: $1 - \frac{1}{\delta^2}$

Pdf: $\frac{2}{\delta^3}$

In order to avoid discontinuities,

Pdf $= |X|^3$ for $|X| > a$.

However, the variance of the distribution is not finite since $E(X^2)$ is not finite.

Since it needs integral with logarithm behavior

Hence, there can be no continous distribution over the whole real axis where the chebyshev Bound is tight.

(Citation is behind)

由 扫描全能王 扫描创建

2. (a) $B(1,d,4) = \{ (x_1 \dots x_d) \in \mathbb{R}^d \mid |x_1^4 + \dots x_d^4| \le 1 \}$

$S: \{x_1 \dots x_d\} \in \mathbb{R}^d \mid |x_1^4 + \dots + x_d^4| \le \frac{1}{2} \}$

Since $vol(B(r,d,4)) = r^d \times vol(B(1,d,4))$

$$\frac{vol(S)}{Vol(B(1,d,4))} = \frac{(\frac{1}{2})^{\frac{1}{4}d} \times Vol(B(1,d,4))}{Vol(B(1,d,4))}$$

$$= (\frac{1}{2})^{\frac{1}{4} \cdot d}$$

(b) $T = \{ x \in \mathbb{R}^d \mid |x| \le 1, x_1 > \varepsilon \}$

$|x_1| < \frac{c}{d^{\frac{1}{4}}}$

$$Vol(T) = \int_{x_1=\varepsilon}^{x_1=1} (1 - x_1^4)^{\frac{d-1}{4}} \cdot vol(B(1, d-1, 4)) dx$$

$$\le (const) \int_{x_1=\varepsilon}^{x_1=1} \frac{x_1^3}{\varepsilon^3} \cdot e^{-\frac{(d-1) \cdot x_1^4}{4}}$$

Since $\int x_1^3 \cdot e^{\frac{-(d-1)x_1^4}{4}} dx = \frac{-1}{d-1} e^{-\frac{(d-1)x_1^4}{4}}$

$$\le (\frac{const}{\varepsilon^3}) \cdot \int_{x_1=\varepsilon}^{x_1=1} x_1^3 \cdot e^{-\frac{(d-1) \cdot x_1^4}{4}} dx$$

$$= (const) \cdot \frac{1}{\varepsilon^3 (d-1)} \cdot e^{-\frac{(d-1)\varepsilon^4}{4}}$$

$$ratio \le \frac{(const) \cdot \frac{1}{\varepsilon^3 \cdot (d-1)} e^{\frac{-(d-1) \cdot \varepsilon^4}{4}}}{(const) \cdot \frac{1}{(d-1)^{\frac{1}{4}}}}$$

$$\le \frac{1}{\varepsilon^3 \cdot (d-1)^{\frac{3}{4}}} e^{-\frac{(d-1) \cdot \varepsilon^4}{4}}$$

let $\varepsilon = \frac{c}{(d-1)^{\frac{1}{4}}}$

$$ratio \le \frac{1}{c^3} e^{-\frac{c^4}{4}}$$

Hence, the fraction of the volume of $B(1,d,4)$ outside the slab $|x_1| \le \frac{c}{d^{\frac{1}{4}}}$ is at most $\frac{1}{c^3} \cdot e^{-\frac{c^4}{4}}$

由 扫描全能王 扫描创建

3. (a) Unit sphere $\Rightarrow$

$B(1,d) = \{x_1 \dots x_d \in \mathbb{R}^d \mid x_1^2 + \dots x_d^2 = 1\}$

Since $E(X_i) = 0$ $\quad E(X) = 0$

(b) $Var(X_i) = E(X_i^2) - E(X_i)^2$

$= \frac{1}{d} E(\sum_{i=1}^{d} X_i^2) - E(X_i)^2$

$= \frac{1}{d}$

(c). $Var(U^T X) = U^2 \cdot var(X)$

Base case: when $d = 1$

$var(u_1^T x_1) = u_1^2 \, var(X_1)$

$= u_1^2 (E(X_1^2) - E(X_1)^2)$

$= u_1^2 \cdot E(X_1^2)$

Suppose $Var(U_k^T X_k) = \sum_{i=1}^{k} u_i^2 \cdot E\{X_i^2\}$

When $d = k+1$,

$Var(u_{k+1}^T x_{k+1}) = Var(U_k^T X_k) + u_{k+1}^2 \cdot E(X_{k+1}^2)$

By induction. $= \sum_{i=1}^{k} u_i^2 \cdot E(X_i^2) + u_{k+1}^2 \cdot E(X_{k+1}^2)$

$= \sum_{i=1}^{k+1} u_i^2 \cdot E(X_i^2)$

Since $Var(X_i) = E(X_i^2) = \frac{1}{d}$.

$Var(u^T X) = \frac{1}{d} \sum_{i=1}^{d} u_i^2 = \frac{1}{d}$

Standard deviation $= \sqrt{\frac{1}{d} \sum_{i=1}^{d} u_i^2} = \sqrt{\frac{1}{d}}$

3. (d) Since ratio $\leq \frac{2}{\varepsilon \cdot \sqrt{d-1}} \cdot e^{-\frac{(d-1)\varepsilon^2}{2}}$

where $\varepsilon = \frac{\alpha}{2}$

ratio $\leq \frac{4}{\alpha \cdot \sqrt{d-1}} \cdot e^{-\frac{(d-1)\alpha^2}{8}}$

$< \frac{8}{\alpha \cdot \sqrt{d-1}} \cdot e^{-\frac{(d-1)\alpha^2}{8}}$

Hence, the volume of their insection is at most $\frac{8 e^{-\alpha^2 \cdot (d-1)/8}}{\alpha \sqrt{d-1}}$

4. In order to make the distance between their images Ax and Ay hugely distorted.
Let $x \neq y$ and $Ax = Ay$
For $A \in \mathbb{R}^{k \times d}$ with $k < d$
let x, y in the nullspace of A.
then $Ax = Ay = 0$
Hence, for every fixed dimension reduction Matrix $A \in \mathbb{R}^{k \times d}$ with $k < d$, there is a pair of vectors $x, y \in \mathbb{R}^d$, such that the distance between their images Ax and Ay is hugely distorted.

3. (e) Since $c \gg 1$ and $r = 1$
let $\alpha = \frac{c}{\sqrt{d-1}}$, then, the fraction of the intersection is at most
$\frac{8 e^{-\frac{c^2}{8}}}{c}$, which is very small mass compared with c when $c \gg 1$

Random variables for which Markov, Chebyshev inequalities are tight. (n.d.). Retrieved March 01, 2018, from https://stats.stackexchange.com/questions/235524/random-variables-for-which-markov-chebyshev-inequalities-are-tight