

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC MỎ-ĐỊA CHẤT

BÁO CÁO HỌC THUẬT

CHẶN CHERNOFF VÀ CHẶN Hoeffding

ThS Nguyễn Thu Hằng

Hà nội, tháng 6 năm 2023

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC MỎ-ĐỊA CHẤT

BÁO CÁO HỌC THUẬT

CHẶN CHERNOFF VÀ CHẶN Hoeffding

Xác nhận của bộ môn

Hà Nội, tháng 6 năm 2023

MỤC LỤC

Lời giới thiệu	1
1. Kiến thức chuẩn bị	2
2. Chặn Chernoff	3
3. Một số ví dụ về ứng dụng của chặn Chernoff	9

LỜI GIỚI THIỆU

Bất đẳng thức là một vấn đề quan trọng của toán học. Nó là một dạng toán tương đối khó và thường không có một phương pháp nào dành riêng để giải quyết một loạt các bài toán này. Mỗi một học phần toán trong chương trình đại học, sinh viên cũng sẽ được tiếp xúc với những bất đẳng thức khác nhau. Trong môn học xác suất thống kê cũng vậy. Có rất nhiều bất đẳng thức xác suất rất thú vị và có nhiều ứng dụng trong lý thuyết cũng như thực tế. Trong báo cáo ở học kỳ I năm học 2022-2023, tôi đã trình bày một số bất đẳng thức xác suất quen thuộc, gần gũi cùng một số ví dụ thực tiễn có thể áp dụng trong giảng dạy. Theo luồng nghiên cứu đó, với mong muốn làm phong phú thêm bài giảng, trong nội dung báo cáo này, tôi tiếp tục nghiên cứu về chặn Chernoff và cùng một số ứng dụng của chặn Chernoff. Từ những bất đẳng thức này, chúng ta có được những ước lượng xác suất tốt hơn.

Báo cáo học thuật chia làm ba phần

Phần 1: Trình bày về hàm sinh mô-men và nhắc lại bất đẳng thức Markov và Chebyshev.

Phần 2: Trình bày về chặn Chernoff.

Phần 3. Trình bày một số ví dụ cơ bản có thể ứng dụng chặn Chernoff.

1. Kiến thức chuẩn bị

1.1. Hàm sinh mômen

Định nghĩa. Cho X là một biến ngẫu nhiên bất kỳ. Hàm sinh mô-men của X là

$$M_X(t) = E[e^{tX}].$$

Nếu X là biến ngẫu nhiên rời rạc

$$M_X(t) = \sum_{i=1}^{\infty} e^{tx_i} p_i.$$

Nếu X là biến ngẫu nhiên liên tục

$$M_X(t) = \int_{-\infty}^{+\infty} e^{tx} f(x) dx.$$

Tính chất của hàm sinh mô-men

i) Khai triển $e^{tX} = 1 + tX + \frac{t^2 X^2}{2!} + \dots + \frac{t^n X^n}{n!} + \dots$

Suy ra

$$M_X(t) = 1 + tEX + \frac{t^2 E(X^2)}{2!} + \dots + \frac{t^n E(X^n)}{n!} + \dots$$

Do đó, hàm sinh mô-men bao gồm tất cả các mô-men của X .

ii) Nếu X_1, X_2, \dots, X_n là các biến ngẫu nhiên độc lập và $X = X_1 + X_2 + \dots + X_n$ thì

$$M_X(t) = \prod_{i=1}^n M_{X_i}(t).$$

Thật vậy, vì X_1, X_2, \dots, X_n độc lập nên $e^{tX_1}, e^{tX_2}, \dots, e^{tX_n}$ cũng độc lập. Ta có

$$\begin{aligned} M_X(t) &= E[e^{tX}] = E[e^{t(X_1 + X_2 + \dots + X_n)}] \\ &= E\left[\prod_{i=1}^n e^{tX_i}\right] = \prod_{i=1}^n E[e^{tX_i}] = \prod_{i=1}^n M_{X_i}(t). \end{aligned}$$

Hàm sinh mô-men của một số phân phối thường gặp

Phân phối	Hàm sinh mô-men
Phân phối 0-1 tham số p	$1 - p + pe^t$
Phân phối nhị thức $B(n, p)$	$(1 - p + pe^t)^n$
Phân phối Poisson	$e^{\lambda(e^t - 1)}$
Phân phối đều liên tục $U(a, b)$	$\frac{e^{tb} - e^{ta}}{t(b - a)}$
Phân phối mũ $E(\lambda)$	$(1 - t\lambda^{-1})^{-1}, t < \lambda$
Phân phối chuẩn $N(\mu, \sigma^2)$	$e^{t\mu + \frac{1}{2}\sigma^2 t^2}$

1.2. Một số bất đẳng thức xác suất khác

Bất đẳng thức Markov. Cho X là một biến ngẫu nhiên nhận giá trị không âm. Khi đó, với mọi $a \geq 0$ ta có

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

Bất đẳng thức Chebyshev. Với mọi $a > 0$ ta có

$$P(|X - E[X]| \geq a) \leq \frac{Var[X]}{a^2}.$$

2. Chặn Chernoff

Chặn Chernoff cho một biến ngẫu nhiên X nhận được bằng cách áp dụng bất đẳng thức Markov đối với e^{tX} cho một số giá trị của t . Cụ thể là:

Với mọi $t > 0$ ta có

$$P(X \geq a) = P(e^{tX} \geq e^{ta}) \leq \frac{E[e^{tX}]}{e^{ta}} \quad (1).$$

Suy ra

$$P(X \geq a) \leq \min_{t>0} \frac{E[e^{tX}]}{e^{ta}}.$$

Với $t < 0$, cũng từ bất đẳng thức Markov suy ra

$$P(X \leq a) = P(e^{tX} \geq e^{ta}) \leq \frac{E[e^{tX}]}{e^{ta}}.$$

Do đó,

$$P(X \leq a) \leq \min_{t > 0} e^{ta} E[e^{-tX}].$$

Như vậy, chìa khóa trong chặn Chernoff chính là hàm sinh mô-men.

Tiếp theo, chúng ta quan tâm đến tổng của n biến ngẫu nhiên độc lập

$$X = X_1 + X_2 + \cdots + X_n.$$

Nếu áp dụng bất đẳng thức Chebychev, với $t > 0$ ta thu được

$$P\left(\left|\sum_{i=1}^n X_i - \sum_{i=1}^n E[X_i]\right| \geq t\right) \leq \frac{\sum_{i=1}^n \text{Var}[X_i]}{t^2}$$

Nếu X_1, X_2, \dots, X_n độc lập cùng phân phối, cộng thêm $E[X_i] = \mu, \forall i = 1, 2, \dots, n$ thì

$$P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| \geq t\right) \leq \frac{\text{Var}[X_1]}{nt^2}$$

Ta sẽ sử dụng chặn Chernoff trong những trường hợp đặc biệt để tìm ước lượng xác suất tốt hơn.

Định lý 1. Cho $X_i, i = 1, 2, \dots, n$ là các biến ngẫu nhiên độc lập, nhận giá trị trong đoạn $[0, 1]$. Đặt $X = X_1 + X_2 + \cdots + X_n$ và giả sử $E[X] = \mu$. Khi đó, với mọi $a > 0$ ta thu được

- i) $P(X \geq \mu + a) \leq e^{-a^2/2n},$
- ii) $P(X \leq \mu - a) \leq e^{-a^2/2n}.$

Chứng minh.

Đặt $Y_k = X_k - E[X_k]$ thì $E[Y_k] = 0, k = 1, 2, \dots, n$.

Đặt $Y = Y_1 + \cdots + Y_n$ thì $Y = X - \mu$. Áp dụng (1), với $t > 0$, ta có

$$P(X \geq \mu + a) = P(Y \geq a) = P(e^{tY} \geq e^{ta}) \leq \frac{E[e^{tY}]}{e^{ta}}.$$

Suy ra

$$P(X \geq \mu + a) \leq \frac{E[e^{tY}]}{e^{ta}} = \frac{E[e^{t \sum_{k=1}^n Y_k}]}{e^{ta}}.$$

Vì các biến ngẫu nhiên X_k độc lập nên $Y_k, k = 1, \dots, n$ là độc lập. Theo tính chất của hàm sinh mô-men, ta có

$$P(X \geq \mu + a) \leq \frac{E[\prod_{i=1}^n e^{tY_k}]}{e^{ta}} = e^{-at} \prod_{k=1}^n E[e^{tY_k}].$$

Tiếp theo ta tìm chặn cho $E[e^{tY_k}]$. Xét hàm số $f(y) = e^{ty}$ có $f''(y) = t^2 e^{ty} > 0$. Gọi $c + dy$ là đường thẳng đi qua hai điểm $(-1, e^{-t})$ và $(1, e^t)$. Ta có thể dễ dàng tính được

$$c = \frac{e^t + e^{-t}}{2}; d = \frac{e^t - e^{-t}}{2}.$$

Dễ thấy

$$e^{ty} = f(y) \leq c + dy, \forall y \in [-1, 1].$$

Suy ra

$$E[e^{tY_k}] \leq E[c + dY_k] = c + dE[Y_k] = c = \frac{e^t + e^{-t}}{2}.$$

Mặt khác, áp dụng khai triển Maclaurin ta được

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$$

Và

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \dots$$

Do đó

$$\begin{aligned} E[e^{tY_k}] &\leq \frac{e^t + e^{-t}}{2} = 1 + \frac{t^2}{2!} + \frac{t^4}{4!} + \frac{t^6}{6!} + \dots \\ &\leq 1 + \frac{t^2}{2 \cdot 1!} + \frac{t^4}{2^2 \cdot 2!} + \frac{t^6}{2^3 \cdot 3!} + \dots \text{ (do } (2k)! \geq 2^k k!). \\ &= e^{t^2/2}. \end{aligned}$$

Ta được

$$P(X \geq \mu + a) \leq e^{-at} \prod_{k=1}^n E[e^{tY_k}] \leq e^{-at} \prod_{k=1}^n e^{\frac{t^2}{2}} = e^{\frac{nt^2}{2} - at}.$$

Hàm số $h(t) = \frac{nt^2}{2} - at$ đạt giá trị nhỏ nhất tại $t_0 = \frac{a}{n}$. Do đó,

$$P(X \geq \mu + a) \leq e^{\frac{n(\frac{a}{n})^2}{2} - a \cdot \frac{a}{n}} = e^{-\frac{a^2}{2n}}.$$

Kết quả ii) thu được bằng cách xét $X' = -X$ và chú ý rằng $X \leq \mu - a$ khi và chỉ khi $X' \geq -\mu + a$. Áp dụng i) ta được

$$P(X \leq \mu - a) \leq e^{-\frac{a^2}{2n}}.$$

Đối với X_k là dãy biến ngẫu nhiên độc lập, có phân phối Bernoulli (phân phối 0-1) thì ta có kết quả sau đây:

Định lý 2. Giả sử $X = X_1 + X_2 + \dots + X_n$ là tổng của n biến ngẫu nhiên Bernoulli độc lập với $P(X_k = 1) = p_k, k = 1, 2, \dots, n$. Đặt $\mu = E[X] = p_1 + \dots + p_n$. Khi đó, với mọi $\delta > 0$, ta có

$$\begin{aligned} \text{i)} \quad & P(X \geq (1 + \delta)\mu) \leq e^{\delta\mu}(1 + \delta)^{-(1+\delta)\mu}, \\ \text{ii)} \quad & P(X \leq (1 - \delta)\mu) \leq e^{\delta\mu}(1 - \delta)^{-(1-\delta)\mu}. \end{aligned}$$

Chứng minh.

Với $t > 1$, ta có

$$P(X \geq (1 + \delta)\mu) = P(t^X > t^{(1+\delta)\mu}).$$

Theo bất đẳng thức Markov, ta được

$$\begin{aligned} P(t^X > t^{(1+\delta)\mu}) &\leq \frac{E[t^X]}{t^{(1+\delta)\mu}} = \frac{E[t^{X_1} t^{X_2} \dots t^{X_n}]}{t^{(1+\delta)\mu}} = \frac{E[t^{X_1}] \dots E[t^{X_n}]}{t^{(1+\delta)\mu}} \\ &= \frac{\prod_{k=1}^n (1 - p_k + tp_k)}{t^{(1+\delta)\mu}} = \frac{\prod_{k=1}^n (1 + p_k(t - 1))}{t^{(1+\delta)\mu}} \\ &\leq \frac{\prod_{k=1}^n e^{p_k(t-1)}}{t^{(1+\delta)\mu}} = \frac{e^{(t-1)(p_1+p_2+\dots+p_n)}}{t^{(1+\delta)\mu}} = \frac{e^{(t-1)\mu}}{t^{(1+\delta)\mu}}. \end{aligned}$$

Suy ra

$$P(X \geq (1 + \delta)\mu) \leq \frac{e^{(t-1)\mu}}{t^{(1+\delta)\mu}}.$$

Thay $t = 1 + \delta$ thì $\delta > 0$. Ta thu được

$$P(X \geq (1 + \delta)\mu) \leq \frac{e^{\delta\mu}}{(1 + \delta)^{(1+\delta)\mu}}.$$

Kết quả ii) được chứng minh tương tự.

Định lí trên cho ta chặn tốt, tuy nhiên ta không thường dùng những kết quả này. Để thuận lợi hơn, ta có thể dùng hệ quả sau đây

Hệ quả 1. Giả sử $X = X_1 + X_2 + \dots + X_n$ là tổng của n biến ngẫu nhiên Bernoulli độc lập với $P(X_k = 1) = p_k, k = 1, 2, \dots, n$. Đặt $\mu = E[X] = p_1 + \dots + p_n$. Khi đó,

$$\text{i)} \quad P(X \geq (1 + \delta)\mu) \leq e^{-\frac{\delta^2\mu}{3}} \text{ với } 0 < \delta < 1,$$

$$\text{ii)} \quad P(X \leq (1 - \delta)\mu) \leq e^{-\frac{\delta^2\mu}{2}} \text{ với } 0 < \delta < 1.$$

Chứng minh.

Để chứng minh i) ta sẽ chỉ ra

$$e^\delta (1 + \delta)^{-(1+\delta)} \leq e^{-\frac{\delta^2}{3}} \text{ với } 0 < \delta < 1. \quad (*)$$

Thật vậy, lấy ln 2 vế, bất đẳng thức tương đương

$$f(\delta) := \delta - (1 + \delta) \ln(1 + \delta) + \frac{\delta^2}{3} \leq 0.$$

Ta có

$$f'(\delta) = 1 - \frac{1 + \delta}{1 + \delta} - \ln(1 + \delta) + \frac{2\delta}{3} = -\ln(1 + \delta) + \frac{2\delta}{3}.$$

Dễ thấy, $f'(\delta) \leq 0$ với mọi $\delta \in [0, 1]$. Do đó, $f(\delta)$ là hàm giảm trên $[0, 1]$ và $f(0) = 0$.

Suy ra

$$f(\delta) = -(1 + \delta) \ln(1 + \delta) + \frac{\delta^2}{3} \leq 0,$$

(*) được chứng minh. Suy ra

$$P(X \geq (1 + \delta)\mu) \leq e^{-\frac{\delta^2\mu}{3}} \text{ với } 0 < \delta < 1.$$

Bất đẳng thức ii) được chứng minh tương tự.

Hệ quả 2. Giả sử $X = X_1 + X_2 + \dots + X_n$ là tổng của n biến ngẫu nhiên Bernoulli độc lập với $P(X_k = 1) = p_k, k = 1, 2, \dots, n$. Đặt $\mu = E[X] = p_1 + \dots + p_n$. Khi đó, với $0 < \delta < 1$ ta có

$$P(|X - \mu| \geq \delta\mu) \leq 2e^{\frac{-\delta^2\mu}{3}}.$$

Hệ quả 3. Giả sử $X = X_1 + X_2 + \dots + X_n$ là tổng của n biến ngẫu nhiên Bernoulli độc lập với $P(X_k = 1) = p, k = 1, 2, \dots, n$. Đặt $\bar{X} = \frac{X}{n}$. Khi đó, với mọi $\delta \in (0, 1), \theta < p$ và $n > \frac{3}{\theta^2} \ln\left(\frac{2}{\delta}\right)$ thì

$$P(|\bar{X} - p| \leq \theta) \geq 1 - \delta.$$

Chứng minh.

Ta có $EX = E[X_1 + X_2 + \dots + X_n] = np$.

Áp dụng hệ quả 2 ta có

$$P(|\bar{X} - p| \geq \varepsilon p) = P(|X - np| \geq \varepsilon np) \leq 2e^{\frac{-\varepsilon^2 np}{3}}.$$

Chọn $\varepsilon = \frac{\theta}{p}$ thì $\theta = \varepsilon p$, ta được

$$P(|\bar{X} - p| \geq \theta) \leq 2e^{\frac{-n\theta^2}{3p}} \leq 2e^{\frac{-n\theta^2}{3}}.$$

Để có $P(|\bar{X} - p| \leq \theta) \geq 1 - \delta$ ta cần $P(|\bar{X} - p| \geq \theta) \leq \delta$. Do đó phải có

$$2e^{\frac{-\varepsilon^2 np}{3}} \leq \delta \Leftrightarrow n > \frac{3}{\theta^2} \ln\left(\frac{2}{\delta}\right).$$

Hệ quả 4. Giả sử $X = X_1 + X_2 + \dots + X_n$ là tổng của n biến ngẫu nhiên Bernoulli độc lập với $P(X_k = 1) = p_k, k = 1, 2, \dots, n$. Đặt $\mu = E[X] = p_1 + \dots + p_n$. Nếu $\theta \geq 6\mu$ thì $P(X \geq \theta) \leq 2^{-\theta}$.

Chứng minh.

Ta đặt $\theta = (1 + \delta)\mu$. Để $\theta \geq 6\mu$ thì chỉ cần lấy $\delta \geq 5$. Áp dụng Định lí 2 ta được

$$P(X \geq (1 + \delta)\mu) \leq \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}}\right)^\mu \leq \left(\frac{e}{1 + \delta}\right)^{(1+\delta)\mu} \leq \left(\frac{e}{6}\right)^\theta \leq 2^{-\theta}.$$

3. Một số ví dụ về ứng dụng của chặn Chernoff

Ví dụ 1. Thực hiện gieo một đồng xu cân đối đồng chất 10000 lần. Giả sử $X_k = 1$ nếu lần gieo thứ k xuất hiện mặt ngửa. Ta có $P(X_k = 1) = \frac{1}{2}$. Đặt $X = X_1 + X_2 + \dots + X_{10000}$. Khi đó

$$E[X] = E[X_1] + \dots + E[X_{10000}] = \frac{10000}{2} = 5000,$$

$$Var[X] = Var[X_1] + \dots + Var[X_n] = \frac{10000}{4} = 2500.$$

Nếu áp dụng bất đẳng thức Markov ta được

$$P(X \geq 6000) \leq \frac{5000}{6000} = \frac{5}{6}.$$

Nếu áp dụng bất đẳng thức Chebychev ta được

$$P(X \geq 6000) = P(X - EX \geq 1000) \leq \frac{2500}{10^6}.$$

Nếu áp dụng chặn Chernoff ta được

$$P(X \geq 6000) = P(X \geq EX + 1000) \leq e^{-\frac{10^6}{2 \cdot 10000}} = e^{-50}.$$

Rõ ràng chặn Chernoff cho ta ước lượng tốt nhất.

Ví dụ 2. Rất nhiều thuật toán ngẫu nhiên có thể mô tả qua trò chơi ném bóng vào thùng sau đây:

Ném m quả bóng đều vào n thùng một cách ngẫu nhiên. Chúng ta quan tâm đến sự phân bố của số bóng trong các thùng. Một số câu hỏi được chú ý là

1. Có bao nhiêu thùng rỗng?
2. Chúng ta phải ném bao nhiêu quả bóng để tất cả các thùng không rỗng?
3. Số bóng tối đa của một thùng là bao nhiêu

Với $i = 1, 2, \dots, n$, giả sử I_i là biến ngẫu nhiên nhận giá trị bằng 1 nếu thùng thứ i rỗng và bằng 0 nếu thùng thứ i không rỗng. Đặt $I = I_1 + \dots + I_n$.

Ta có

$$E[I_i] = P(\text{thùng } i \text{ rỗng}) = \prod_{j=1}^m (1 - P(\text{quả } j \text{ vào thùng } i)) = \left(1 - \frac{1}{n}\right)^m$$

Do đó

$$E[I] = E[I_1] + \dots + E[I_n] = n \left(1 - \frac{1}{n}\right)^m.$$

Như vậy

$$P(\text{tồn tại thùng rỗng}) \leq \sum_{i=1}^n P(\text{thùng } i \text{ rỗng}) \leq n \left(1 - \frac{1}{n}\right)^m \leq ne^{-\frac{m}{n}}.$$

Trong trường hợp $m \approx n$, thì $ne^{-\frac{m}{n}} = \frac{n}{e} > 1$. Do đó, đánh giá là không có tác dụng. Để có thể sử dụng đánh giá này, ta cần giả thiết $m \gg n \ln n$. Trong trường hợp này ta thấy rằng, xác suất để tồn tại thùng rỗng là rất nhỏ.

Với $i = 1, 2, \dots, n$, giả sử X_i là biến ngẫu nhiên chỉ số quả bóng rơi vào thùng thứ i . Đặt $Z = \max\{X_1, \dots, X_n\}$.

Gọi X_i^j là biến ngẫu nhiên nhận giá trị 1 nếu quả thứ j rơi vào hộp i và bằng 0 trong trường hợp ngược lại. Ta có $E[X_i^j] = \frac{1}{n}$ và do đó $E[X_i] = \frac{m}{n}$. Áp dụng Định lý 2, ta được

$$P\left(X_i \geq (1 + \delta) \frac{m}{n}\right) \leq \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}}\right)^{\frac{m}{n}}.$$

Chú ý rằng nếu $\delta > e - 1$ thì

$$\frac{e^\delta}{(1 + \delta)^{1+\delta}} = \frac{1}{1 + \delta} \left(\frac{e}{1 + \delta}\right)^\delta < \frac{1}{1 + \delta} < \frac{1}{2}.$$

Do đó, nếu $m > 2n \log_2 n$ thì ta được

$$P\left(X_i \geq (1 + \delta) \frac{m}{n}\right) \leq 2^{-\frac{m}{n}} \leq \frac{1}{n^2}.$$

Và ta được

$$P\left(Z \geq (1 + \delta) \frac{m}{n}\right) \leq P(Z \geq e \cdot \frac{m}{n}) \leq n \cdot n^{-2} = \frac{1}{n}.$$

Ta cũng có

$$P\left(X_i \geq (1 + \delta) \frac{m}{n}\right) \leq \left(\frac{e}{1 + \delta}\right)^{(1+\delta)\frac{m}{n}}.$$

Do vậy, nếu $m = n$, chọn $\delta = \frac{c \ln n}{\ln \ln n} - 1$ với c đủ lớn thì

$$P(X_i \geq 1 + \delta) \leq n^{-2}.$$

Khi đó

$$P(Z \geq 1 + \delta) \leq n \cdot n^{-2} = \frac{1}{n}.$$

Ví dụ 3. Trong lí thuyết độ phức tạp, người ta quan tâm đến BPP (**bounded-error probabilistic polynomial**) là lớp các bài toán quyết định giải được bằng máy Turing ngẫu nhiên trong thời gian đa thức, với xác suất sai không quá $1/3$ cho mọi trường hợp. Cho hằng số $\varepsilon \in [0, \frac{1}{3})$. Một ngôn ngữ L nằm trong lớp BPP nếu tồn tại thuật toán ngẫu nhiên với thời gian đa thức A thỏa mãn

1. Nếu $x \in L$ suy ra $P(A(x) \text{ chấp nhận}) \geq 1 - \varepsilon$,
2. Nếu $x \notin L$ suy ra $P(A(x) \text{ từ chối}) \geq 1 - \varepsilon$.

Giả sử A là một thuật toán ngẫu nhiên quyết định $L \in \text{BPP}$. Ta xây dựng thuật toán A' chạy A trên đầu vào x với số lần là n . Vậy A' mắc lỗi với xác suất ra sao?

Kí hiệu X_k là biến ngẫu nhiên bằng 1 nếu lần chạy thứ k mà trả về kết quả đúng. Đặt $X = X_1 + \dots + X_n$ thì X chỉ số lần trả về kết quả đúng trong n lần chạy. Ta có

$$E[X] \geq \frac{2}{3}n.$$

Do đó,

$$P\left(\sum_{k=1}^n X_k \leq \frac{n}{2}\right) \leq P\left(\sum_{k=1}^n X_k \leq EX - \frac{n}{6}\right) \leq e^{-\frac{(\frac{n}{6})^2}{2n}} = e^{-\frac{n}{72}}.$$

Nếu chọn $n = 72 \ln \frac{1}{\delta}$ thì thuật toán sai số với xác suất δ .

KẾT LUẬN

Báo cáo đã trình bày về chặn Chernoff với các dạng thể hiện khác nhau và một số ví dụ thú vị có thể sử dụng chặn Chernoff. Tài liệu có thể dùng làm tài liệu tham khảo cho giảng viên, sinh viên, làm sinh động thêm những ứng dụng của môn xác suất thống kê đối với các vấn đề thực tiễn trong cuộc sống. Tuy nhiên, đây chỉ là một nội dung nhỏ trong vô vàn kết quả về bất đẳng thức xác suất. Trong thời gian tới, tác giả sẽ tiếp tục tìm hiểu về chặn Hoeffding cũng như nhiều bất đẳng thức khác nhằm hoàn thiện hơn về nội dung nghiên cứu và nâng cao chất lượng bài giảng cũng như làm tốt mục tiêu gắn lý thuyết với thực tiễn.

TÀI LIỆU THAM KHẢO

1. N.D.Tiến, V.V.Yên, *Lý thuyết xác suất*, NXB Giáo dục 2006.
2. N.C.Văn, T.T.Ninh, *Giáo trình Lý thuyết xác suất và thống kê toán*, NXB Đại học Kinh tế quốc dân, 2004
3. N.T.T.Hà, *Một số bất đẳng thức xác suất và ứng dụng*, Khóa luận tốt nghiệp Đại học.