

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

BÁO CÁO HỌC THUẬT

MỘT SỐ KHOẢNG CÁCH
GIỮA CÁC BIẾN NGẪU NHIÊN
(Phần 2)

ThS. Nguyễn Thu Hằng

HÀ NỘI, 6/2026

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

BÁO CÁO HỌC THUẬT

MỘT SỐ KHOẢNG CÁCH
GIỮA CÁC BIẾN NGẪU NHIÊN
(Phần 2)

Xác nhận của Bộ môn Toán

HÀ NỘI, 6/2026

Mục lục

1	Định nghĩa và tính chất của lượng thông tin Fisher và khoảng cách lượng thông tin Fisher	4
1.1	Định nghĩa và tính chất của lượng thông tin Fisher	4
1.2	Định nghĩa và tính chất của khoảng cách lượng thông tin Fisher	9
2	Mối liên hệ của khoảng cách lượng thông tin Fisher với các khoảng cách khác	10

MỞ ĐẦU

Trong Lý thuyết xác suất và thống kê toán học, hàm mật độ và hàm phân phối tích lũy (gọi chung là luật xác suất) là hai khái niệm cơ bản nhất của một biến ngẫu nhiên. Tuy nhiên trong hầu hết các mô hình ứng dụng thực tiễn, chẳng hạn như nghiệm của một phương trình vi phân ngẫu nhiên (Stochastic Differential Equation viết tắt là SDE), hai hàm này thường không thể tìm được công thức hiển. Do đó, việc nghiên cứu luật xác suất của một biến ngẫu nhiên là rất quan trọng và đã được các nhà khoa học quan tâm từ nhiều thế kỷ. Thông thường, có hai cách nghiên cứu luật xác suất của biến ngẫu nhiên, thứ nhất là nghiên cứu trực tiếp, thứ hai là nghiên cứu gián tiếp thông qua ước lượng mức độ gần, đo sự khác biệt hay khoảng cách giữa chúng với một biến ngẫu nhiên quen thuộc nào đó. Đây là bài toán quan trọng của lý thuyết xác suất và thống kê toán học vì nó không chỉ cho chúng ta sự so sánh giữa các phân phối, giữa mô hình và hiện thực mà còn mang lại nhiều thông tin về sự hội tụ, ước tính sai số, độ rủi ro và tối ưu hóa trong thuật toán. Có rất nhiều khoảng cách khác nhau được sử dụng như khoảng cách Kolmogorov, khoảng cách Wasserstein, khoảng cách biến phân toàn phần, khoảng cách thông tin Fisher...

Liên quan đến lượng thông tin Fisher (Fisher information) đã có một lịch sử nghiên cứu lâu dài, bắt nguồn từ công trình của Linnik năm 1959 [?]. Phần lớn các kết quả hiện có trong lĩnh vực này tập trung vào tổng của các biến ngẫu nhiên độc lập. Đối với các tổng này, nhiều kết quả định lượng về tốc độ hội tụ của lượng thông tin Fisher đã được thiết lập, tiêu biểu là các công trình của Bobkov, Johnson và cộng sự [1, 2, 3, 4]. Cùng với đó là nhiều công trình nghiên cứu khoảng cách lượng thông tin Fisher $I(F||N)$ giữa biến ngẫu nhiên F và biến ngẫu nhiên có phân phối chuẩn N [5]. Chú ý rằng, $I(F||N)$ không phải là một khoảng cách thực sự. Tuy nhiên đại lượng này liên quan chặt chẽ đến nhiều loại khoảng cách quan trọng khác trong thống kê ứng dụng, chẳng hạn như entropy tương đối (relative entropy hay còn gọi là khoảng cách Kullback–Leibler) và khoảng cách supremum giữa các mật độ xác suất. Ngoài ra, $I(F||N)$ cũng có mối liên hệ với các khoảng cách truyền thống như khoảng cách Kolmogorov, khoảng cách biến phân toàn phần và khoảng cách Wasserstein, v.v. Do đó, $I(F||N)$ đóng vai trò như một thước đo mạnh đo lường sự khác biệt giữa

quy luật phân bố của một biến ngẫu nhiên so với phân phối chuẩn dựa vào cấu trúc vi phân của hàm mật độ xác suất và có nhiều ứng dụng trong nghiên cứu Định lý giới hạn trung tâm.

Trong báo cáo này, chúng tôi tìm hiểu và trình bày một số kiến thức cơ bản về lượng thông tin Fisher và khoảng cách lượng thông tin Fisher giữa biến ngẫu nhiên F và phân phối chuẩn Gauss cũng như liên hệ giữa khoảng cách lượng thông tin Fisher với một số khoảng cách thường gặp khác.

Nội dung báo cáo gồm hai phần:

- Phần 1 trình bày tóm tắt một số kiến thức cơ bản về lượng thông tin Fisher và khoảng cách lượng thông tin Fisher như định nghĩa, một số tính chất và các ví dụ tính toán cho một số phân phối thường gặp.
- Phần 2 trình bày mối liên hệ giữa khoảng cách lượng thông tin Fisher với một số khoảng cách khác như L^1 —khoảng cách, khoảng cách Kolmogorov, khoảng cách biến phân toàn phần.

NỘI DUNG BÁO CÁO

1 Định nghĩa và tính chất của lượng thông tin Fisher và khoảng cách lượng thông tin Fisher

1.1 Định nghĩa và tính chất của lượng thông tin Fisher

Định nghĩa 1.1. Cho F là biến ngẫu nhiên có hàm mật độ liên tục tuyệt đối p_F . Hàm score ρ_F của biến ngẫu nhiên F được định nghĩa bởi

$$\rho_F := \frac{p'_F(x)}{p_F(x)}$$

và lượng thông tin Fisher $I(F)$ của biến ngẫu nhiên F được định nghĩa bởi

$$I(F) := \int_{-\infty}^{+\infty} \frac{p'_F(x)^2}{p_F(x)} dx.$$

Từ định nghĩa, ta dễ dàng suy ra một số tính chất sau:

1. $I(F) = E[\rho_F^2(F)]$.
2. $E[\rho_F(F)] = 0$.
3. $I(F) \geq 0$.
4. $I(aX + b) = \frac{1}{a^2} I(X)$, $\forall a, b \in \mathbb{R}, a \neq 0$.

Ví dụ 1. Nếu N là biến ngẫu nhiên có phân phối chuẩn với trung bình μ và phương sai σ^2 thì

- Hàm mật độ của N là $\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.
- Hàm score của N là $\rho_N = \frac{-(x-\mu)}{\sigma^2}$.
- Lượng thông tin Fisher của N là

$$\begin{aligned} I(N) &= E[\rho_N^2(N)] = \frac{1}{\sigma^4} \int_{-\infty}^{+\infty} (x-\mu)^2 \varphi(x) dx \\ &= \frac{1}{\sigma^4} \int_{-\infty}^{+\infty} (x^2 - 2\mu x + \mu^2) \varphi(x) dx \\ &= \frac{1}{\sigma^4} (E[N^2] - 2\mu E[N] + \mu^2) = \frac{1}{\sigma^4} \text{Var}[N] = \frac{1}{\sigma^2}. \end{aligned}$$

Ví dụ 2. Nếu F là biến ngẫu nhiên có phân phối mũ với tham số λ thì

- Hàm mật độ của F là $f(x) = \lambda e^{-\lambda x}, x \geq 0$.
- Hàm score của F là $\rho_F = -\lambda$.
- Lượng thông tin Fisher của F là $I(F) = E[\rho_F^2] = \lambda^2$.

Ví dụ 3. Nếu F là biến ngẫu nhiên có phân phối Gamma $F \sim \Gamma(n, \theta)$ thì

- Hàm mật độ của F là $f(x) = \frac{e^{-\theta x} \theta^n x^{n-1}}{\Gamma(n)}$.
- Hàm score của F là $\rho_F = -\theta + \frac{n-1}{x}$. Sử dụng tính chất $EF^s = \frac{\Gamma(n+s)}{\Gamma(n)\theta^s}$ nếu $s > -n$ ta được
- Lượng thông tin Fisher của F là

$$\begin{aligned} I(F) = E[\rho_F^2] &= (n-1)^2 \frac{\theta^2}{(n-2)(n-1)} - 2(n-1)\theta \frac{\theta}{n-1} + \theta^2 \\ &= \frac{\theta^2}{n-1}, n > 2. \end{aligned}$$

Mệnh đề 1.2. Cho N cho biến ngẫu nhiên có phân phối chuẩn với trung bình μ và phương sai σ^2 . Nếu F là biến ngẫu nhiên có trung bình μ và phương sai σ^2 với hàm mật độ liên tục tuyệt đối ρ_F thỏa mãn $\lim_{x \rightarrow \pm\infty} (x - \mu)p_F(x) = 0$ thì

$$I(F) \geq I(N).$$

Dấu bằng xảy ra khi và chỉ khi $F \sim N(\mu, \sigma^2)$.

Chứng minh. Giả sử F có hàm mật độ p_F . Với f là hàm khả vi thỏa mãn $\lim_{x \rightarrow \pm\infty} f(x)p_F(x) = 0$ ta có

$$\begin{aligned} E[f(F)\rho(F)] &= \int_{-\infty}^{+\infty} f(x) \frac{p'_F(x)}{p_F(x)} p_F(x) dx = \int_{-\infty}^{+\infty} p'_F(x) f(x) dx \\ &= - \int_{-\infty}^{+\infty} f'(x) p_F(x) dx = -E[f'(F)]. \end{aligned}$$

Áp dụng cho $f(x) = ax + b$ ta có $E[(aF + b)\rho_F(F)] = -a$. Do đó

$$\begin{aligned} 0 &\leq E \left[\left(\rho_F(F) + \frac{F - \mu}{\sigma^2} \right)^2 \right] \\ &= I(F) + 2E \left[\frac{F - \mu}{\sigma^2} \rho_F(F) \right] + E \left[\frac{F - \mu}{\sigma^2} \right]^2 \\ &= I(F) - \frac{2}{\sigma^2} + \frac{1}{\sigma^2} = I(F) - I(N). \end{aligned}$$

Vậy $I(F) \geq I(N)$.

Dấu bằng xảy ra khi và chỉ khi

$$\frac{p'_F(x)}{p_F(x)} = \frac{d \ln(p_F(x))}{dx} = -\frac{x - \mu}{\sigma^2} \quad h.k.n.$$

Nghĩa là

$$p_F(x) = C e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad h.k.n.,$$

trong đó C là hằng số thỏa mãn các điều kiện để $p_F(x)$ là hàm mật độ. Nói cách khác, dấu bằng xảy ra khi và chỉ khi F có phân phối chuẩn. \square

Mệnh đề 1.3. Nếu X, Y là các biến ngẫu nhiên độc lập thì với mọi $\beta \in [0, 1]$

$$I(X + Y) \leq \beta^2 I(X) + (1 - \beta)^2 I(Y) \quad (1)$$

và

$$I(\sqrt{\beta}X + \sqrt{1 - \beta}Y) \leq \beta I(X) + (1 - \beta)I(Y). \quad (2)$$

Chứng minh. Đặt $W = X + Y$. Giả sử hàm mật độ của X, Y lần lượt là f, g . Khi đó, hàm mật độ của W là

$$p_W(w) = \int f(x)g(w - x)dx.$$

Suy ra

$$p'_W(w) = \int f(x) \frac{\partial g}{\partial w}(w - x)dx = - \int f(x) \frac{\partial g}{\partial x}(w - x)dx = \int f'(x)g(w - x)dx.$$

Do đó, hàm score của W là

$$\begin{aligned} \rho_W(w) = \frac{p'_W(w)}{p_W(w)} &= \int \frac{f'(x)g(w - x)}{p_W(w)}dx = \int \frac{f'(x)}{f(x)} \frac{f(x)g(w - x)}{p_W(w)}dx \\ &= E \left[\frac{f'(X)}{f(X)} \middle| W = w \right] = E[\rho_X(X) | W = w]. \end{aligned}$$

Tương tự, $\rho_W(w) = E[\rho_Y(Y)|W = w]$. Suy ra, với mọi $\beta \in [0, 1]$ ta có

$$\rho_W(w) = E[\beta\rho_X(X) + (1 - \beta)\rho_Y(Y)|W = w]. \quad (3)$$

Áp dụng bất đẳng thức Jensen ta thu được

$$\begin{aligned} I(W) &= E[\rho_W^2(W)] = E\left[(E[\beta\rho_X(X) + (1 - \beta)\rho_Y(Y)|W = w])^2\right] \\ &\leq E\left[E(\beta\rho_X(X) + (1 - \beta)\rho_Y(Y))^2|W = w\right] \\ &= \beta^2 E[\rho_X^2(X)] + (1 - \beta)^2 E[\rho_Y^2(Y)]. \end{aligned}$$

Trong ước lượng (1), thay $\sqrt{\beta}U$ và $\sqrt{1 - \beta}V$ cho U và V tương ứng và chú ý rằng $I(aX) = \frac{1}{a^2}I(X)$ ta thu được (2). \square

Thông tin Fisher cũng có thể định nghĩa theo tham số, dùng để đo lường thông tin của biến ngẫu nhiên X về tham số θ mà xác suất của X phụ thuộc vào nó.

Định nghĩa 1.4. Gọi $f(X, \theta)$ là hàm mật độ của X (trong trường hợp X là biến ngẫu nhiên liên tục) và là hàm khối lượng của X (mass function trong trường hợp X là biến ngẫu nhiên rời rạc) thì lượng thông tin Fisher $I_\theta(X)$ về tham số θ của biến ngẫu nhiên X được định nghĩa bởi

$$I_\theta(X) = E\left[\left(\frac{\partial}{\partial\theta} \log f(X, \theta)\right)^2\right].$$

Ví dụ 4. Nếu biến ngẫu nhiên có phân phối 0-1 $X \sim \text{Bernoulli}(p)$ thì

$$f(x, p) = P(X = x) = p^x(1 - p)^{1-x}, \quad x \in \{0, 1\}$$

và $EX = EX^2 = p$. Khi đó

$$\log f(x, p) = x \log p + (1 - x) \log(1 - p).$$

Lấy đạo hàm theo p ta được

$$\frac{\partial}{\partial\theta} \log f(X, p) = \frac{x}{p} - \frac{1 - x}{1 - p}.$$

Suy ra

$$\begin{aligned}
 I_p(X) &= E \left[\left(\frac{X}{p} - \frac{1-X}{1-p} \right)^2 \right] \\
 &= E \left[\frac{X^2}{p^2} - \frac{2F(1-X)}{p(1-p)} + \frac{(1-X)^2}{(1-p)^2} \right] \\
 &= \frac{p}{p^2} + \frac{2(p-p)}{p(1-p)} + \frac{1-p}{(1-p)^2} \\
 &= \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)}.
 \end{aligned}$$

Ví dụ 5. Nếu biến ngẫu nhiên X có phân phối Poisson với tham số λ thì

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

và $EX = VarX = \lambda E[X^2] = \lambda + \lambda^2$.

Khi đó

$$\log f(x, \lambda) = -\lambda + x \log \lambda - \log(x!).$$

Lấy đạo hàm theo λ ta được

$$\frac{\partial}{\partial \theta} \log f(x, \lambda) = -1 + \frac{x}{\lambda}.$$

Suy ra

$$\begin{aligned}
 I_\lambda(X) &= E \left[\left(-1 + \frac{X}{\lambda} \right)^2 \right] \\
 &= E \left[1 - \frac{2X}{\lambda} + \frac{X^2}{\lambda^2} \right] \\
 &= 1 - \frac{2\lambda}{\lambda} + \frac{\lambda + \lambda^2}{\lambda^2} = \frac{1}{\lambda}.
 \end{aligned}$$

Ví dụ 6. Nếu biến ngẫu nhiên X có phân phối mũ với tham số λ thì hàm mật độ

$$f(x, \lambda) = \lambda e^{-\lambda x}, \quad x \geq 0.$$

và $EX = \frac{1}{\lambda}$, $VarX = \frac{1}{\lambda^2}$, $E[X^2] = \frac{2}{\lambda^2}$. Khi đó

$$\log f(x, \lambda) = \log \lambda - \lambda x.$$

Lấy đạo hàm theo λ ta được

$$\frac{\partial}{\partial \theta} \log f(x, \lambda) = \frac{1}{\lambda} - x.$$

Suy ra

$$\begin{aligned}
I_\lambda(X) &= E \left[\left(\frac{1}{\lambda} - X \right)^2 \right] \\
&= E \left[\frac{1}{\lambda^2} - 2\frac{X}{\lambda} + X^2 \right] \\
&= \frac{1}{\lambda^2} - \frac{2}{\lambda^2} + \frac{2}{\lambda^2} = \frac{1}{\lambda^2}.
\end{aligned}$$

1.2 Định nghĩa và tính chất của khoảng cách lượng thông tin Fisher

Định nghĩa 1.5. Cho F là biến ngẫu nhiên với hàm mật độ liên tục tuyệt đối. Khoảng cách lượng thông tin Fisher của F đến N được định nghĩa bởi

$$I(F\|N) := E \left[\left(\rho_F(F) + \frac{F - \mu}{\sigma^2} \right)^2 \right].$$

Chú ý rằng, $I(F\|N)$ không phải là một khoảng cách thực sự. Nếu F có trung bình μ , phương sai σ^2 và hàm mật độ ρ_F thỏa mãn $\lim_{x \rightarrow \pm\infty} (x - \mu)p_F(x) = 0$ thì tương tự trong chứng minh của Mệnh đề 1.2 ta có

$$I(F\|N) = I(F) - I(N).$$

Ví dụ 7. Giả sử F là biến ngẫu nhiên có phân phối mũ với tham số λ thì $E[F] = \frac{1}{\lambda}$ và $Var[F] = \frac{1}{\lambda^2}$. Giả sử N là biến ngẫu nhiên chuẩn có kì vọng và phương sai bằng với kì vọng và phương sai của F . Khi đó

$$\rho_F(F) + \frac{F - \mu}{\sigma^2} = -\lambda + \lambda^2 \left(F - \frac{1}{\lambda} \right) = \lambda^2 F - 2\lambda.$$

Suy ra

$$\begin{aligned}
I(F\|N) &= E[(\lambda^2 F - 2\lambda)^2] \\
&= \lambda^4 E[F^2] - 4\lambda^3 E[F] - 4\lambda^2 \\
&= \lambda^4 \left(Var[F] + (E[F])^2 \right) - 4\lambda^3 E[F] - 4\lambda^2 \\
&= \lambda^4 \frac{2}{\lambda^2} - 4\lambda^3 \frac{1}{\lambda} - 4\lambda^2 \\
&= 2\lambda^2.
\end{aligned}$$

2 Mối liên hệ của khoảng cách lượng thông tin Fisher với các khoảng cách khác

Tiếp theo, ta tìm hiểu mối quan hệ giữa khoảng cách lượng thông tin Fisher và một số khoảng cách khác thường dùng.

Định nghĩa 2.1. Xét không gian đo (Ω, \mathcal{F}) và P, Q là hai độ đo trên (Ω, \mathcal{F}) . Khoảng cách biến phân toàn phần (d_{TV}) giữa P, Q được định nghĩa bởi

$$d_{TV}(P, Q) = \sup_{A \in \mathcal{F}} |P(A) - Q(A)|.$$

Nói cách khác, nếu X, Y là hai biến ngẫu nhiên thì

$$d_{TV}(X, Y) = \sup_{D \in \mathcal{B}(\mathbb{R})} |P(X \in D) - P(Y \in D)|.$$

Định nghĩa 2.2. Khoảng cách Kolmogorov giữa hai biến ngẫu nhiên X, Y được định nghĩa bởi

$$d_K(X, Y) := \sup_{x \in \mathbb{R}} |P(X \leq x) - P(Y \leq x)|.$$

Định nghĩa 2.3. Cho X, Y là hai biến ngẫu nhiên liên tục có hàm mật độ lần lượt là $p(x)$ và $q(x)$. Khi đó khoảng cách entropy (độ lệch entropy tương đối) từ X tới Y cho bởi

$$D_{KL}(X \| Y) = \int_{\mathbb{R}} p(x) \log \left(\frac{p(x)}{q(x)} \right) dx.$$

Do các tập $(-\infty, x]$ chỉ là một lớp con của mọi tập đo được nên

$$d_K(X, Y) \leq d_{TV}(X, Y).$$

Mệnh đề 2.4 (Bất đẳng thức Csiszár-Kullback-Pinsker).

$$d_{TV}(X, Y) \leq \sqrt{\frac{1}{2} D_{KL}(X \| Y)}. \quad (4)$$

Chứng minh. Báo cáo trình bày chứng minh của Clément L. Canonne trong [7].

Xét X và Y là các phân phối 0 – 1 với tham số lần lượt là p, q . Theo ví dụ 1 của mục 1 ta có

$$d_{TV}(X, Y) = |p - q|$$

và

$$D_{KL}(X\|Y) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}.$$

Xét hàm $f(x) : (0, 1) \rightarrow \mathbb{R}$ cho bởi $f(x) = p \log x + (1-p) \log(1-x)$ thì ta có thể viết

$$\begin{aligned} f(p) - f(q) &= \int_q^p f'(x) dx = \int_q^p \frac{p-x}{x(1-x)} dx \\ &\leq 4 \int_q^p (p-x) dx = 4 \cdot \frac{1}{2} (p-q)^2. \end{aligned}$$

Ở trên, ta đã sử dụng bất đẳng thức $x(1-x) \leq 1/4$ với $x \in (0, 1)$. Từ đó suy ra

$$f(p) - f(q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q} \leq 2(p-q)^2.$$

Nên bất đẳng thức 4 đúng trong trường hợp này. Trong trường hợp tổng quát, giả sử X, Y là biến ngẫu nhiên với trên miền Ω tùy ý tương ứng với độ đo xác suất P, Q trên Ω . Cố định tập con đo được $S \subset \Omega$. Hai biến ngẫu nhiên $X' = \mathbb{1}_S(X)$ và $Y' = \mathbb{1}_S(Y)$ tương ứng là các phân phối 0-1 với tham số $p' = P(S)$ và $q' = Q(S)$. Do đó

$$2(P(S) - Q(S))^2 = 2d_{TV}(X', Y')^2 \leq D_{KL}(X', Y') \leq D_{KL}(X\|Y),$$

trong đó bất đẳng thức thứ 2 là nhờ bất đẳng thức data processing. Lấy sup trên các tập S ta được

$$2d_{TV}(X, Y)^2 = 2 \sup_{S \in \Omega} 2(P(S) - Q(S))^2 \leq D_{KL}(X\|Y).$$

□

Mệnh đề 2.5. Cho F là biến ngẫu nhiên có trung bình μ , phương sai σ^2 và hàm mật độ liên tục tuyệt đối. Khi đó, ta có

$$d_{TV}(F, N) \leq \frac{\sigma}{2} \sqrt{I(F\|N)}. \quad (5)$$

Chứng minh. entropy tương đối (relative entropy) của F đối với N là $D_{KL}(F\|N)$, cho bởi

$$D_{KL}(F\|N) = \int_{\mathbb{R}} p_F(x) \log \left(\frac{p_F(x)}{p_N(x)} \right) dx.$$

Theo bất đẳng thức Csiszár-Kullback-Pinsker ta có

$$2(d_{TV}(F, N))^2 \leq D(F\|N).$$

Mặt khác, theo Bổ đề 1 trong [6], ta có

$$D_{KL}(F\|N) = \int_0^1 \frac{J(\sqrt{t}F + \sqrt{1-t}N)}{2t} dt,$$

với N là biến ngẫu nhiên độc lập với F . Áp dụng bất đẳng thức (2), suy ra

$$D_{KL}(F\|N) \leq \int_0^1 \frac{tJ(F) + (1-t)J(N)}{2t} dt = \frac{\sigma^2}{2} I(F\|N). \quad (6)$$

Kết hợp các ước lượng ta được

$$(d_{TV}(F, N))^2 \leq \frac{\sigma^2}{4} I(F\|N).$$

Lấy căn thức hai vế ta được điều phải chứng minh □

Bất đẳng thức (6) còn được biết đến là bất đẳng thức log-Sobolev.

Như vậy, khoảng cách lượng thông tin Fisher là một khoảng cách mạnh. Sự hội tụ của khoảng cách lượng thông tin Fisher sẽ dẫn đến sự hội tụ của của một số khoảng cách khác như khoảng cách biến phân toàn phần, khoảng cách Kolmogorov hay khoảng cách entropy.

KẾT LUẬN

Báo cáo đã tổng kết một số kiến thức cơ bản về lượng thông tin Fisher và khoảng cách khoảng cách lượng thông tin Fisher giữa các biến ngẫu nhiên. Từ đó, tính toán lượng thông tin Fisher và khoảng cách khoảng cách lượng thông tin Fisher giữa một số biến ngẫu nhiên cơ bản. Báo cáo cũng tổng hợp một số mối quan hệ giữa khoảng cách khoảng cách lượng thông tin Fisher và một số khoảng cách thường dùng khác. Kết quả của báo cáo nhằm mục đích làm tài liệu tham khảo chuyên ngành xác suất thống kê và hỗ trợ giảng dạy cho sinh viên ngành toán.

TÀI LIỆU THAM KHẢO

- [1] S. G. Bobkov, G. P. Chistyakov, F. Götze, Berry-Esseen bounds in the entropic central limit theorem. *Probab. Theory Related Fields* 159 (2014), no. 3-4, 435–478.
- [2] O. Johnson, A. R. Barron, Fisher information inequalities and the central limit theorem. *Probab. Theory Related Fields*, 129 (2004), no. 3, 391–409.
- [3] O. Johnson, *Information theory and the central limit theorem*. Imperial College Press, London, 2004.
- [4] O. Johnson, Maximal correlation and the rate of Fisher information convergence in the central limit theorem. *IEEE Trans. Inform. Theory* 66 (2020), no. 8, 4992–5002.
- [5] I. Nourdin, D. Nualart, Fisher information and the fourth moment theorem. *Ann. Inst. Henri Poincaré Probab. Stat.* 52 (2016), no. 2, 849–867.
- [6] Andrew R. Barron, Entropy and the Central Limit Theorem, *The Annals of Probability*, Vol. 14, No. 1 (Jan., 1986), pp. 336-342.
- [7] Clément L. Canonne. A short note on an inequality between KL and TV, <https://arxiv.org/abs/2202.07198v2>