

به نام خدا

آنتروپی تفاضلی

آنتروپی تفاضلی مشترک

آنتروپی تفاضلی مشترک برای n متغیر تصادفی $x_1, x_2, \dots, x_n$ تابع حتمی احتمال

توان $f_x(x_1, x_2, \dots, x_n)$ به صورت زیر تعریف می شود،

$$h(x_1, x_2, \dots, x_n) \equiv h(f) \equiv h(x^n) = - \int f(\underline{x}) \log f(\underline{x}) d\underline{x}$$

آنتروپی توافقی شرط

اگر x, y در متغیرهای تابع احتمالات $f(x, y)$ باشند، آنتروپی
توافقی برای آن به صورت زیر تعریف می شود:

$$h(x|y) = - \int f(x, y) \lg f(x, y) \, dx \, dy$$

$$\underbrace{f(x, y) / f(y)} \quad \text{یعنی} \quad f(x, y) = \frac{f(x, y)}{f(y)}$$

با توجه به اینکه

$$h(x|y) = h(x, y) - h(y)$$

فصل: آلر $x_1, x_2, \dots, x_n$ متغیرهای مستقل دارای توزیع با مترس کارایی κ باشند، آنتروپی تعاملی مشترک برای آنها از رابطه زیر بدست می آید:

$$h(x^n) \equiv h(\underline{x}) \equiv \frac{1}{2} \log(2\pi e |\kappa|)$$

در بردار κ بسایس n

$$|\kappa| \equiv |\det \kappa|$$

$$f_x(\underline{x}) = \varphi_x(\underline{x}) = \frac{1}{\sqrt{(2\pi)^n \det \kappa}} \exp\left(-\frac{1}{2} \underline{(x-\mu)}^T \kappa^{-1} (\underline{x}-\mu)\right)$$

اصت κ این است که آنتروپی از صحت κ nat κ کنیم

$$E\left\{-\frac{1}{2} (\underline{x}-\mu)^T \kappa^{-1} (\underline{x}-\mu)\right\} = \frac{1}{2} E\left\{\text{trace}((\underline{x}-\mu)^T \kappa^{-1} (\underline{x}-\mu))\right\} =$$

(تمرین)

$$= -\frac{1}{2} \in \left\{ \text{trace} \left(\underline{(x-\mu)(x-\mu)^T} K^{-1} \right) \right\} = \dots$$

آنتروپی نسبی یافته K

برای در تابع چگالی احتمال $f_x(x)$ ، آنتروپی نسبی یافته K به صورت زیر تعریف می شود:

$$D(f||g) = \int f \log \frac{f}{g}$$

اصول و متقابل

برای در معرقتاری x و y با تابع چگالی احتمال تمام $f(x, y)$ در تابع احتمال ناری $f_x(x)$ و $f_y(y)$ ، اصول و متقابل به صورت زیر تعریف می شود:

$$I(x; y) = \int f(x, y) \log \frac{f(x, y)}{f(x)f(y)} dx dy$$

$$\equiv D(f(x, y) \| f(x)f(y))$$

$$= h(x) - h(x|y) = h(y) - h(y|x) = h(x) + h(y) - h(x, y)$$

مثال: در شفره صافرخس x, y ترمای نهال با بیشترین صفر و مایوس کارایس

$$K = \begin{bmatrix} \epsilon^2 & \rho\epsilon^2 \\ \rho\epsilon^2 & \epsilon^2 \end{bmatrix}$$

پسند. $I(x; y) = ?$

$$L(x; y) = h(x) + h(y) - h(x, y)$$

$$h(x) = \frac{1}{2} \lg(2\pi e \zeta^2) = h(y)$$

$$h(x, y) = \frac{1}{2} \lg(2\pi e)^2 |k| = \frac{1}{2} \lg(2\pi e)^2 \zeta^4 (1 - \rho^2)$$

$$\Rightarrow L(x; y) = \frac{1}{2} \lg(2\pi e \zeta^2) + \frac{1}{2} \lg(2\pi e \zeta^2) - \frac{1}{2} \lg(2\pi e)^2 \zeta^4 (1 - \rho^2)$$

$$= \frac{1}{2} \lg \frac{(2\pi e)^2 \zeta^4}{(2\pi e)^2 \zeta^4 (1 - \rho^2)} = \frac{1}{2} \lg \left(\frac{1}{1 - \rho^2} \right)$$

$$X \perp\!\!\!\perp Y \rightarrow \rho = 0 \Rightarrow I(X; Y) = 0$$

$$|P| = 1 \rightarrow \text{رابطه خطی بین } x, y \text{ برقرار است} \Rightarrow I(X; Y) \rightarrow \infty$$

نقشه: اکثریتی نسبتاً صلی $D(f \| g) \geq 0$ و متساوی برقرار است اگر و تنها اگر $f \stackrel{a.e.}{=} g$ w.p. 1

نتیجه: ۱- $I(X; Y) \geq 0$ و متساوی برقرار است اگر و تنها اگر $X \perp\!\!\!\perp Y$

۲- $h(X|Y) \leq h(X)$ و متساوی برقرار است اگر و تنها اگر $X \perp\!\!\!\perp Y$

* قانون زنجیره‌ای

$$h(x_1, x_2, \dots, x_n) = \sum_{i=1}^n h(x_i | x_{i-1}, \dots, x_1)$$

$$h(x_1, x_2, \dots, x_n) \leq \sum_{i=1}^n h(x_i)$$

تسبیح:

ستادی برقرار است اگر و تنها اگر
 $x_1 \perp x_2 \perp \dots \perp x_n$

* ناسیادی حاد دارد

اگر $x_1, x_2, \dots, x_n$ تواباً نرمال با ماتریس کوارانس K باشند، همواره داریم
برای ساینین صفر
 $|K| \leq \prod_{i=1}^n k_{ii}$

ساده برقرار است اگر داشته آید $x_1 \perp x_2 \perp \dots \perp x_n$

اثبات:

$$h(x_1, x_2, \dots, x_n) \leq \sum_{i=1}^n h(x_i)$$

$$\Rightarrow \frac{1}{2} \log (2\pi e)^n |K| \leq \sum_{i=1}^n \frac{1}{2} \log (2\pi e \underbrace{\sigma_i^2}_{k_{ii}})$$

$$\Rightarrow \frac{1}{n} \log (2\pi e)^n |K| \leq \frac{1}{n} \log \prod_{i=1}^n (2\pi e k_{ii})$$

$$\Rightarrow \log |K| \leq \log \prod_{i=1}^n k_{ii} \quad \Rightarrow |K| \leq \prod_{i=1}^n k_{ii}$$

قضیه: شیب دادن متغیرها در انتگرال تغییر نمی دهد یعنی

$$h(x+c) = h(x)$$

$$-\int_{x+c} f \log f \equiv -\int_x f \log f$$

اثبات:

$$d(x+c) = dx$$

$$h(\overset{y}{ax}) = h(x) + \log |a|$$

قضیه:

$$f(\overset{y}{ax}) = \frac{1}{|a|} f_x\left(\frac{y}{a}\right)$$

اثبات:

$$h(y) = -\int f_y(y) \log f_y(y) dy = -\int \frac{1}{|a|} f_x\left(\frac{y}{a}\right) \log\left(\frac{1}{|a|} f_x\left(\frac{y}{a}\right)\right) dy$$

$$\Rightarrow h(y) = - \int \frac{1}{|a|} f_x\left(\frac{y}{a}\right) \left[\log \frac{1}{|a|} + \log f_x\left(\frac{y}{a}\right) \right] dy$$

$$= - \underbrace{\int \frac{1}{|a|} f_x\left(\frac{y}{a}\right) \log \frac{1}{|a|} dy}_{\log \frac{1}{|a|}} - \int \frac{1}{|a|} f_x\left(\frac{y}{a}\right) \log f_x\left(\frac{y}{a}\right) dy$$

$\underbrace{\quad}_{adn}$

$$= \log |a| + h(x)$$

$$h(A\underline{x}) = h(\underline{x}) + \log |\det A|$$

↑
تبدیل از سبب شغف

تقسیم به حالت عادی :

قضیه: اگر $\underline{x} \in \mathbb{R}^n$ و $x_1, x_2, \dots, x_n$ دارای میانگین صفر و ماتریس کوارانس

$$\Sigma \text{ باشد، داریم}$$
$$h(\underline{x}) \leq \frac{1}{2} (2\pi e)^n |\Sigma|$$

ستادی وقتی برقرار است که $x_1, x_2, \dots, x_n$ توانا نرمال باشند.

معنی! ماتریس کوارانس ثابت، بردار نرمال ماکزیم آنبردنی را به دست می دهد.

اثبات: فرض کنیم که بردار $\underline{x}$ دارای تابع چگالی احتمال $g(\underline{x})$ باشد، ماتریس کوارانس
دین بردار ستادین برابر Σ باشد. توزیع $\varphi_k(\underline{x})$ یک توزیع نرمال به صورت $N(0, \Sigma)$
در نظر می گیریم. می دانیم که همواره

$$D(g \parallel \varphi_k) \geq 0$$

$$g \stackrel{a.e}{=} \varphi_k$$

دستاری برقرار است اگر دانه اگر
از حُرُفِ رَسْمِ

$$D(g \| \varphi_k) = \int g \log \frac{g}{\varphi_k}$$

$$= \int g \log g - \int g \log \varphi_k \geq 0$$

$$= - \frac{h(x)}{h(g)} - \int g \log \varphi_k$$

$x^T K^{-1} x$ است.

نرم‌های φ_k به صورت

$$\int g \log \varphi_k = \int g (x^T K^{-1} x) = \int g \left(\sum_i \sum_j d_{ij} x_i x_j \right)$$

$$\kappa^{-1} = [\alpha_{ij}]$$

$$\begin{aligned} \Rightarrow \int g \sum_i \sum_j \alpha_{ij} x_i x_j &\equiv \sum_i \sum_j \alpha_{ij} E_g x_i x_j \\ &= \int \varphi_k \underbrace{\sum_i \sum_j \alpha_{ij} x_i x_j}_{\sum_i \sum_j \alpha_{ij} E_\varphi x_i x_j} \end{aligned}$$

$$\Rightarrow \int g \log \varphi_k \equiv \int \varphi_k \log \varphi_k$$

چون ماتریس کربارمانس در درصت برابر است

$$D(g \| \varphi_k) = -h(g) - \int g \log \varphi_k \geq 0$$

$$\Rightarrow D(g \| \varphi_k) = -h(g) - \int \varphi_k \log \varphi_k = -h(g) + h(\varphi) \geq 0$$

$$\Rightarrow h(g) \leq h(\varphi_k)$$

نقشه: اگر x یک شعریه دهن با اثر دین نامی $h(x)$ باشد، $\hat{x}$ یک تخمین از x ، $E(x - \hat{x})^2$ میانگین خطای مرتبه دوم تخمین باشد، همراه نامگذاری زیر برقرار است.

nat : بر حسب $h(x)$

$$E(x - \hat{x})^2 \geq \frac{1}{2\pi e} e^{2h(x)}$$

ستادین زمان برقرار است که x یک متغیر تصادفی گوسی و $\hat{X} = Ex$ باشد.

اثبات: از تئوری شگفتی می دانیم که برای هر کمیتی دلخواه $\hat{X}$ از متغیر تصادفی x داریم

$$E(x - \hat{X})^2 \geq \min E(x - \hat{X})^2 = E(x - Ex)^2 = \text{Var}(x)$$

$\hat{X} = Ex$ می بینیم برای $\hat{X} = Ex$ حاصل می شود

$$\Rightarrow E(x - \hat{X})^2 \geq \text{Var}(x)$$

از طرف دیگر می دانیم که با درایسن برابر، متغیر تصادفی گوسی، بیشترین آنتروپی را به دست می دهد

$$h(x) \leq h(\varphi_{\text{Var}(x)}) = \frac{1}{2} \log(2\pi e) \text{Var}(x) \Rightarrow \text{Var}(x) \geq \frac{1}{2\pi e} e^{2h(x)}$$

تصادفی متغیری برقرار است که x توزیع کروی داشته باشد و $\hat{x} = \epsilon x$ باشد.

$$\Rightarrow E(x - \hat{x})^2 \geq \text{Var}(x) \geq \frac{1}{2\pi e} e^{2h(x)}$$

$$\Rightarrow E(x - \hat{x})^2 \geq \frac{1}{2\pi e} e^{2h(x)}$$

نتیجه: اگر $\hat{x} = \hat{x}(y)$ بهترین تخمین از متغیر تصادفی x با داشتن مشاهده y باشد،

$$E(x - \hat{x}(y))^2 \geq \frac{1}{2\pi e} e^{2h(x|y)}$$

که نگاه داریم.

بهترین تخمین در این حالت در میان تخمین‌های مرتبه دوم رایج است می‌دهد $\hat{x}(y) = E(x|y)$

بقیه اثبات مثل به حالت قبل است .

سپاس ری وقتی برقرار است که x, y توانا زغال باشند ، $x|y \in \hat{X}$ باشد .

