Informacija ir entropija

Informacijos kiekis

Tarkime A yra tikimybinės erdvės (Ω, P) atsitiktinis įvykis, kurio tikimybė $P(A) = p$. Įykio A prigimtis, skaičiuojant informacijos kiekį $I(A)$, nėra svarbi. Todėl jį apibrėšime kaip kintamojo p funkciją ir dažnai rašysime $I(A) = I(p)$.

Informacijos kiekis

Reikalavimai informacijos kiekiui:

Informacijos kiekis

Reikalavimai informacijos kiekiui:

1. Informacija turi būti apibrėžta ir neneigiama, t.y. $I(p) \geq 0$, visiems $p \in (0,\ 1]$.

Informacijos kiekis

Reikalavimai informacijos kiekiui:

1. Informacija turi būti apibrėžta ir neneigiama, t.y. $I(p) \geq 0$, visiems $p \in (0,\ 1]$.
2. Nežymiai pakitus įvykio tikimybei, informacijos kiekis taip pat turėtų pasikeisti nedaug. Kitaip sakant funkcija $I(p)$ turi būti tolydi.

Informacijos kiekis

Reikalavimai informacijos kiekiui:

1. Informacija turi būti apibrėžta ir neneigiama, t.y. $I(p) \geq 0$, visiems $p \in (0,\ 1]$.
2. Nežymiai pakitus įvykio tikimybei, informacijos kiekis taip pat turėtų pasikeisti nedaug. Kitaip sakant funkcija $I(p)$ turi būti tolydi.
3. Funkcija $I(p)$ turi būti griežtai monotoniškai mažėjanti, t.y. kuo įvykio tikimybė mažesnė, tuo didesnį informacijos kiekį jam įvykus gauname.

Informacijos kiekis

Reikalavimai informacijos kiekiui:

1. Informacija turi būti apibrėžta ir neneigiama, t.y. $I(p) \geq 0$, visiems $p \in (0,\ 1]$.
2. Nežymiai pakitus įvykio tikimybei, informacijos kiekis taip pat turėtų pasikeisti nedaug. Kitaip sakant funkcija $I(p)$ turi būti tolydi.
3. Funkcija $I(p)$ turi būti griežtai monotoniškai mažėjanti, t.y. kuo įvykio tikimybė mažesnė, tuo didesnį informacijos kiekį jam įvykus gauname.
4. Įvykus dviem nepriklausomiems įvykiams, gautos informacijos kiekis turėtų būti lygus jų informacijų sumai. Prisiminę, kad dviems nepriklausomiems įvykiams A ir B tikimybė įvykti kartu yra $P(A \cap B) = P(A)P(B)$, turėsime tokį reikalavimą informacijos kiekio funkcijai: $I(p \cdot q) = I(p) + I(q)$ visiems $p, q \in (0,\ 1]$.

Informacijos kiekio išraiška

Teorema. *Funkcija $I(p)$ tenkina 1-4 sąlygas tada ir tik tada, kai egzistuoja $b > 1$, jog*

$$I(p) = \log_b \frac{1}{p}\,.$$

Įrodymas. Tegul m ir n bet kokie natūralieji skaičiai. Iš 4 sąlygos išplaukia, kad

$$\begin{aligned} I(p^n) &= I(p \cdot p^{n-1}) = I(p) + I(p^{n-1}) \\ &= I(p) + I(p) + I(p^{n-2}) = \cdots = nI(p)\,. \end{aligned}$$

Todėl

$$I(p) = I((p^{1/m})^m) = mI(p^{1/m})$$

ir

$$I(p^{1/m}) = \frac{1}{m} I(p)\,.$$

Informacijos kiekio išraiška

Taigi, visiems teigiamiems racionaliesiems skaičiams $\frac{n}{m}$

$$I(p^{n/m}) = I((p^{1/m})^n) = \frac{n}{m} I(p)\,.$$

Dėl funkcijos $I(p)$ tolydumo iš čia išplaukia, kad

$$I(p^a) = aI(p)$$

visiems realiesiems $a \geq 0$. Todėl visiems $p \in (0,\ 1]$

$$\begin{aligned} I(p) &= I\left(\left(\frac{1}{\mathrm{e}}\right)^{-\ln p}\right) = -I\left(\frac{1}{\mathrm{e}}\right) \ln p \\ &= -\frac{\ln p}{\ln b} = \log_b \frac{1}{p}\,, \quad b > 1. \end{aligned}$$

$\square$

Informacijos kiekio vienetai

Apibrėžimas. *Informacijos kiekiu, gaunamu įvykus įvykiui A, kurio tikimybė $p > 0$, vadinsime dydį*

$$I(A) = I(p) = \log_2 \frac{1}{p} \, ,$$

o jo matavimo vienetus - bitais.

Atskiru atveju $I(1) = 0$, $I(0) = \lim_{p \to 0} I(p) = \infty$.

Kartais naudojami ir kiti informacijos kiekio vienetai.

$Log.\ pagrindas\ (b)$	$Inf.\ kiekio\ vienetas$
2	bitas
3	tritas
e	natas
10	hartlis

$$1\, bitas \;=\log_3 2\, trito \;= \ln 2\, nato \;= \lg 2\, hartlio\, .$$

Pavyzdys. Metame simetrišką monetą. $P(S) = P(H) = 0,5$. Todėl bet kurios baigties atveju gaunamas $I(S) = I(H) = \log_2 2 = 1$ bitas informacijos. Jei moneta metama n kartų, tai bet kurią eksperimento baigtį galime nusakyti n dvejetainių skaitmenų, pavyzdžiui

$$\underbrace{011101110\ldots01101}_{n}$$

Čia 0 ir 1 žymi įvykius S ir H. Tokios baigties tikimybė yra 2^{-n}, o gaunamas informacijos kiekis

$$I\left(\frac{1}{2^n}\right) = \log_2 2^n = n\,,$$

t.y. lygiai tiek, kiek bitų užima informacija apie eksperimento rezultatą.

Sąlyginė informacija

Apibrėžimas. *Tegul A ir B yra tikimybinės erdvės (Ω, P) atsitiktiniai įvykiai ir $P(B) > 0$. Įvykio A su sąlyga B informacija vadinsime dydį*

$$I(A|B) = \log_2 \frac{1}{P(A|B)} = -\log_2 \frac{P(A \cap B)}{P(B)}.$$

Pastebėsime, kad $I(A|B) = I(A)$ tada ir tik tada, kai įvykiai A ir B yra nepriklausomi.

Entropija

Tegul galimos bandymo baigtys yra $A_1, A_2, \dots A_n$, o jų tikimybės atitinkamai $p_1, p_2, \dots p_n$. Atliekame N tokių nepriklausomų bandymų. Tada vidutinis vieno bandymo informacijos kiekis bus

$$\frac{I_N}{N} \approx \frac{1}{N}\sum_{i=1}^{n} N p_i \log_2 \frac{1}{p_i} = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i}\,.$$

Entropija

Tegul galimos bandymo baigtys yra $A_1, A_2, \ldots A_n$, o jų tikimybės atitinkamai $p_1, p_2, \ldots p_n$. Atliekame N tokių nepriklausomų bandymų. Tada vidutinis vieno bandymo informacijos kiekis bus

$$\frac{I_N}{N} \approx \frac{1}{N} \sum_{i=1}^{n} N p_i \log_2 \frac{1}{p_i} = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i} .$$

Apibrėžimas. *Tegul* $\mathcal{A} = \{A_1, A_2, \ldots, A_n\}$ *yra tikimybinės erdvės* (Ω, P) *įvykių su tikimybėmis* $p_i = P(A_i)$ *sistema. Įvykių sistemos* $\mathcal{A}$ *entropija vadinsime dydį*

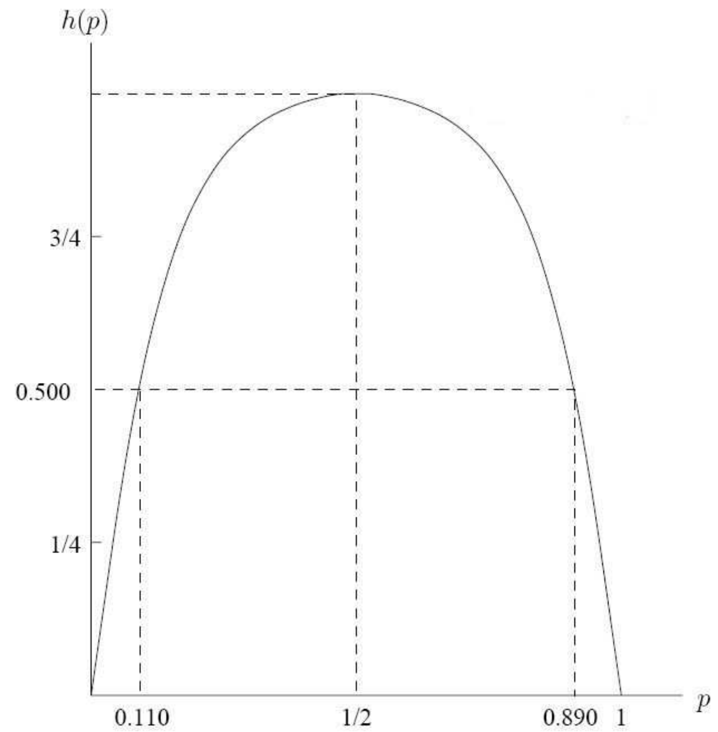
$$H(\mathcal{A}) = H(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i} .$$

Čia ir toliau visi dėmenys $p_i \log_2 \frac{1}{p_i}$ arba $p_i \log_2 p_i$ yra lygūs 0 , kai $p_i = 0$.

Binarinė entropija

Dviejų įvykių su tikimybėmis p ir $1-p$ sistemos entropija

$$h(p) = H(p, 1-p) = p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p} \,.$$



Sąlyginė entropija

Įvykus kokiam nors įvykiui B, sistemos $\mathcal{A}$ entropija gali pasikeisti. Likusio neapibrėžtumo laipsnį nusako sąlyginė entropija

$$H(\mathcal{A}|B) = \sum_{i=1}^{n} P(A_i|B) I(A_i|B) = \sum_{i=1}^{n} P(A_i|B) \log_2 \frac{1}{P(A_i|B)}\,.$$

Tegul $\mathcal{B} = \{B_1, B_2, \ldots, B_m\}$ yra kita tos pačios tikimybinės erdvės įvykių sistema.

Apibrėžimas. *Dydį*

$$H(\mathcal{A}|\mathcal{B}) = \sum_{j=1}^{m} P(B_j) H(\mathcal{A}|B_j)$$

vadinsime sąlygine $\mathcal{A}$ entropija $\mathcal{B}$ atžvilgiu, o entropijos pokytį $I(\mathcal{A}, \mathcal{B}) = H(\mathcal{A}) - H(\mathcal{A}|\mathcal{B})$ vadinsime sistemų $\mathcal{A}$ ir $\mathcal{B}$ tarpusavio informacija.

Jungtinės sistemos entropija

Sistemų $\mathcal{A}$ ir $\mathcal{B}$ jungtinės sistemos $\mathcal{A} \wedge \mathcal{B}$ entropiją žymėsime $H(\mathcal{A}, \mathcal{B})$

$$H(\mathcal{A}, \mathcal{B}) = H(\mathcal{A} \wedge \mathcal{B}) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(A_i \cap B_j) \log_2 \frac{1}{P(A_i \cap B_j)} .$$

Šis apibrėžimas akivaizdžiai apibendrinamas ir didesniam įvykių sistemų skaičiui $k \geq 2$

$$H(\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k) = H(\mathcal{A}_1 \wedge \mathcal{A}_2 \wedge \cdots \wedge \mathcal{A}_k) .$$

Gibbs'o nelygybė

Vektorių $(x_1, x_2, \dots, x_n)$, sudarytą iš neneigiamų realiųjų skaičių, vadinsime tikimybiniu vektoriumi (kitaip: diskrečiuoju skirstiniu), jei

$$\sum_{i=1}^{n} x_i = 1\,.$$

Gibbs'o nelygybė. *Tegul* $b > 1$. *Tada bet kokiems tikimybiniams vektoriams* $(x_1, x_2, \dots, x_n)$ *ir* $(y_1, y_2, \dots, y_n)$ *teisinga nelygybė*

$$\sum_{i=1}^{n} x_i \log_b \left(\frac{y_i}{x_i} \right) \leq 0\,.$$

Nelygybė virsta lygybe tada ir tik tada, kai vektoriai $(x_1, x_2, \dots, x_n)$ *ir* $(y_1, y_2, \dots, y_n)$ *sutampa.*

Gibbs'o nelygybė

Įrodymas. Pakanka įrodyti nelygybę

$$\sum_{i=1}^{n}{}^{*} x_i \ln\left(\frac{y_i}{x_i}\right) \le 0\,.$$

Čia simbolis * prie sumos ženklo reiškia, kad sumuojama tik pagal tas i reikšmes, kurioms $x_i > 0$. Kadangi funkcija $y = \ln x$ yra iškila aukštyn, o jos grafiko liestinė taške $x = 1$ yra tiesė $y = x - 1$, tai $\ln x \le x - 1$ visiems $x > 0$. Be to, nelygybė virsta lygybe tik, kai $x = 1$. Iš čia išplaukia

$$\sum_{i=1}^{n}{}^{*} x_i \ln\left(\frac{y_i}{x_i}\right) \le \sum_{i=1}^{n}{}^{*} x_i \left(\frac{y_i}{x_i} - 1\right) = \sum_{i=1}^{n}{}^{*} y_i - \sum_{i=1}^{n}{}^{*} x_i \le 0\,.$$

Pastebėsime, kad abi pastarosios nelygybės virs lygybėmis tik kai $x_i = y_i$ visiems i. $\square$

Entropijos įverčiai

Teorema. *Įvykių sistemos $\mathcal{A} = \{A_1, A_2, \ldots, A_n\}$ entropija tenkina nelygybes*

$$0 \leq H(\mathcal{A}) \leq \log_2 n\,. \qquad (1)$$

Mažiausią reikšmę $H(\mathcal{A}) = 0$ ji įgyja tada ir tik tada, kai sistema sudaryta iš įvykių, kurių tikimybės yra 0 arba 1. Didžiausią entropiją $H(\mathcal{A}) = \log_2 n$ turės tos ir tik tos sistemos, kuriose visi įvykiai yra vienodai galimi, t.y. $P(A_i) = \frac{1}{n}$ visiems $i = 1, 2, \ldots, n$.

Įrodymas. Pagal apibrėžimą entropija yra neneigiamų dėmenų suma

$$H(\mathcal{A}) = \sum_{i=1}^{n} P(A_i) \log_2 \frac{1}{P(A_i)}\,.$$

Todėl aišku, kad $H(\mathcal{A}) \geq 0$. Be to, tokia suma lygi 0 tada ir tik tada, kai visi dėmenys lygūs 0. Vadinasi, $P(A_i) = 0$ arba $P(A_i) = 1$ visiems $i = 1, 2, \ldots, n$. Iš tikrųjų tik viena iš šių tikimybių bus lygi 1, nes sistemą sudarančių įvykių tikimybių suma visada yra 1.

Entropijos įverčiai

Antrąją nelygybę įrodysime skirtumui $H(\mathcal{A}) - \log_2 n$ pritaikę Gibbs'o nelygybę. Gausime

$$\begin{aligned} H(\mathcal{A}) - \log_2 n &= \sum_{i=1}^{n} P(A_i) \log_2 \frac{1}{P(A_i)} - \sum_{i=1}^{n} P(A_i) \log_2 n \\ &= \sum_{i=1}^{n} P(A_i) \log_2 \left(\frac{1/n}{P(A_i)} \right) \leq 0\,. \end{aligned}$$

$\square$

Pavyzdys

Tegul $\mathcal{A} = \{A_1, A_2, A_3, A_4\}$, $\mathcal{B} = \{B_1, B_2, B_3, B_4, B_5\}$ ir

$$\begin{aligned} P(A_1) &= P(A_2) = P(A_3) = P(A_4) = \frac{1}{4}\,; \\ P(B_1) &= P(B_2) = P(B_3) = P(B_4) = \frac{1}{16}\,, \quad P(B_5) = \frac{3}{4}\,. \end{aligned}$$

Apskaičiuojame sistemų $\mathcal{A}$ ir $\mathcal{B}$ entropijas

$$\begin{aligned} H(\mathcal{A}) &= 4 \cdot \frac{1}{4} \cdot \log_2 4\,; \\ H(\mathcal{B}) &= 4 \cdot \frac{1}{16} \cdot \log_2 16 + \frac{3}{4} \cdot \log_2 \frac{4}{3} \approx 1,3113\,. \end{aligned}$$

Kaip matome, $H(\mathcal{A}) > H(\mathcal{B})$. Gautąją nelygybę galime interpretuoti taip: numatyti, kuris iš įvykių įvyks, sistemoje $\mathcal{A}$ yra sunkiau, nei sistemoje $\mathcal{B}$.

Tikslesnės sistemos entropija

Teorema. *Jei įvykių sistema* $\mathcal{B} = \{B_1, B_2, \ldots, B_m\}$ *yra tikslesnė už sistemą* $\mathcal{A} = \{A_1, A_2, \ldots, A_n\}$ *, tai*

$$H(\mathcal{A}) \leq H(\mathcal{B}).$$

Tikslesnės sistemos entropija

Įrodymas. Kadangi $\mathcal{B}$ yra tikslesnė už $\mathcal{A}$, tai

$$P(A_i) = \sum_{j=1}^{m} P(A_i \cap B_j) = \sum_{j \in J_i} P(B_j)\,, \quad i = 1, 2, \ldots, n.$$

Čia J_i - poromis nesikertantys aibės $\{1, 2, \ldots, m\}$ poaibiai, tenkinantys sąlygą

$$\bigcup_{i=1}^{n} J_i = \{1, 2, \ldots, m\}\,.$$

Tikslesnės sistemos entropija

Todėl

$$
\begin{aligned}
H(\mathcal{A}) &= -\sum_{i=1}^{n} P(A_i)\log_2 P(A_i) \\
&= -\sum_{i=1}^{n}\left(\sum_{j\in J_i} P(B_j)\right)\log_2\left(\sum_{j\in J_i} P(B_j)\right) \\
&\le -\sum_{i=1}^{n}\sum_{j\in J_i} P(B_j)\log_2 P(B_j) \\
&= -\sum_{j=1}^{m} P(B_j)\log_2 P(B_j) = H(\mathcal{B})
\end{aligned}
$$

$\square$

Sąlyginės entropijos pavyzdys

Tegul X ir Y yra atsitiktiniai dydžiai, įgyjantys reikšmes 0 ir 1, o jų bendrasis dvimatis skirstinys nusakytas lentele

$X\backslash Y$	0	1	$P(X=i)$
0	$0,25$	$0,25$	$0,5$
1	0	$0,5$	$0,5$
$P(Y=j)$	$0,25$	$0,75$	1

Nagrinėsime dvi įvykių sistemas

$$\mathcal{A}=\{\,\{X=0\},\ \{X=1\}\,\}\ \text{ ir }\ \mathcal{B}=\{\,\{Y=0\},\ \{Y=1\}\,\}\,.$$

Sistemų $\mathcal{A}$ ir $\mathcal{B}$ entropijas žymėsime tiesiog $H(X)$ ir $H(Y)$. Gausime, kad $H(Y|X=1)<H(Y)<H(Y|X=0)$, tačiau $H(Y|X)<H(Y)$.

Tarpusavio informacijos įvertis

Teorema. *Visoms įvykių sistemoms* $\mathcal{A} = \{A_1, A_2, \dots, A_n\}$ *ir* $\mathcal{B} = \{B_1, B_2, \dots, B_m\}$ *jų tarpusavio informacija yra neneigiama:*

$$I(\mathcal{A}, \mathcal{B}) \geq 0\,.$$

Be to, $I(\mathcal{A}, \mathcal{B}) = 0$ *tada ir tik tada, kai sistemos* $\mathcal{A}$ *ir* $\mathcal{B}$ *yra nepriklausomos.*

Tarpusavio informacijos įvertis

Įrodymas. Pagal apibrėžimą

$$
\begin{aligned}
H(\mathcal{A}|\mathcal{B}) &= \sum_{j=1}^{m} P(B_j) H(\mathcal{A}|B_j) \\
&= \sum_{j=1}^{m} \sum_{i=1}^{n} P(B_j) P(A_i|B_j) \log_2 \frac{1}{P(A_i|B_j)} \\
&= \sum_{j=1}^{m} \sum_{i=1}^{n} P(A_i \cap B_j) \log_2 \left(\frac{P(B_j)}{P(A_i \cap B_j)} \right) .
\end{aligned}
$$

Tarpusavio informacijos įvertis

Iš įvykių sistemos apibrėžimo išplaukia, kad visiems $i = 1, 2, \ldots, n$

$$P(A_i) = \sum_{j=1}^{m} P(A_i \cap B_j)\,.$$

Todėl

$$H(\mathcal{A}) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(A_i \cap B_j) \log_2 \frac{1}{P(A_i)}\,.$$

Tarpusavio informacijos įvertis

Įstatę gautas išraiškas, turėsime

$$\begin{aligned} -I(\mathcal{A},\mathcal{B}) &= H(\mathcal{A}|\mathcal{B}) - H(\mathcal{A}) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} P(A_i \cap B_j)\log_2\left(\frac{P(A_i)P(B_j)}{P(A_i \cap B_j)}\right) \le 0\,. \end{aligned}$$

Pastaroji nelygybė išplaukia iš Gibbs'o nelygybės, pritaikius ją tikimybiniams vektoriams, kurių komponentės yra

$$P(A_i \cap B_j) \text{ ir } P(A_i)P(B_j)\,, \quad i = 1, 2, \ldots, n\,, \ j = 1, 2, \ldots, m\,.$$

Pagal tą pačią lemą gauname, kad nelygybė virsta lygybe tada ir tik tada, kai $P(A_i \cap B_j) = P(A_i)P(B_j)$, visiems i ir j. T.y., kai sistemos $\mathcal{A}$ ir $\mathcal{B}$ yra nepriklausomos.

$\square$

Entropijų sąryšiai

Teorema. *Įvykių sistemoms $\mathcal{A}$ ir $\mathcal{B}$ teisingi sąryšiai*

$$H(\mathcal{A},\mathcal{B}) = H(\mathcal{B}) + H(\mathcal{A}|\mathcal{B})\,, \tag{1}$$

$$H(\mathcal{A},\mathcal{B}) = H(\mathcal{A}) + H(\mathcal{B}) - I(\mathcal{A},\mathcal{B})\,. \tag{2}$$

Įrodymas.

$$\begin{aligned} H(\mathcal{A}|\mathcal{B}) &= \sum_{j=1}^{m}\sum_{i=1}^{n} P(A_i\cap B_j)\log_2\left(\frac{1}{P(A_i\cap B_j)}\right) \\ &- \sum_{j=1}^{m}\left(\sum_{i=1}^{n} P(A_i\cap B_j)\right)\log_2\left(\frac{1}{P(B_j)}\right) \\ &= H(\mathcal{A},\mathcal{B}) - H(\mathcal{B})\,. \end{aligned}$$

(2) gauname iš (1), nes $H(\mathcal{A}|\mathcal{B}) = H(\mathcal{A}) - I(\mathcal{A},\mathcal{B})$. $\square$

Jungtinės sistemos entropijos įvertis

Kadangi $I(\mathcal{A}, \mathcal{B}) \geq 0$, tai

$$H(\mathcal{A}, \mathcal{B}) \leq H(\mathcal{A}) + H(\mathcal{B})\,.$$

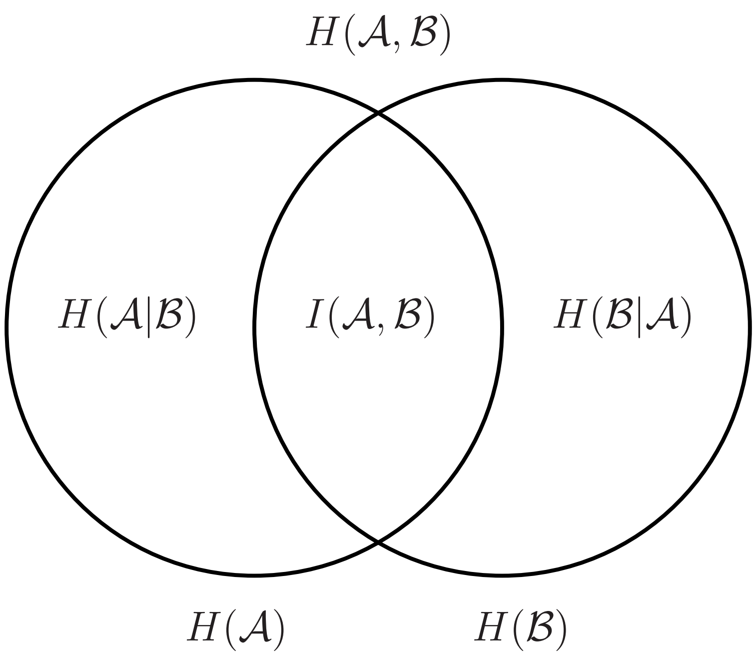
Pritaikius indukciją, pastarąją nelygybę nesudėtinga apibendrinti didesniam įvykių sistemų skaičiui.
Teorema. *Jei $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k$ yra tos pačios diskrečiosios tikimybinės erdvės įvykių sistemos, tai*

$$H(\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k) \leq H(\mathcal{A}_1) + H(\mathcal{A}_2) + \cdots + H(\mathcal{A}_k)\,.$$

Ši nelygybė virsta lygybe tada ir tik tada, kai $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k$ yra nepriklausomos sistemos.

Entropijų sąryšių diagrama

Dviejų sistemų entropijų, sąlyginių entropijų bei tarpusavio informacjos sąryšius vaizduojanti diagrama.



Diskrečiojo a. d. entropija

Apibrėžimas. *Diskrečiojo atsitiktinio dydžio* X, *įgyjančio reikšmes* $x_1, x_2, \ldots$, *entropija* $H(X)$ *yra lygi įvykių sistemos, sudarytos iš įvykių* $A_i = \{X = x_i\}, \ i = 1, 2, \ldots$, *entropijai, t.y.,*

$$H(X) = \sum_i P(A_i) \log_2 \frac{1}{P(A_i)} \, .$$

1pvz. (Šenono rėžis)

Tegul T yra koks nors tekstas. Pavyzdžiui, $T = abrakadabra$
A. d. X reiškia atsitiktinai pasirinktą T raidę.

X	a	b	d	k	r
P	5/11	2/11	1/11	1/11	2/11

$$H(X) = \frac{5}{11}\log_2\frac{11}{5} + 2\cdot\frac{2}{11}\log_2\frac{11}{2} + 2\cdot\frac{1}{11}\log_2 11 \approx 2,0404\,.$$

Tai yra vadinamasis Šenono rėžis vidutiniam vienos raidės kodo bitų skaičiui. Kitaip sakant, nėra tokio vienareikšmiškai dekoduojamo kodo, kuriuo tekstą T atvaizduotume trumpesne nei $11\cdot H(X) \approx 22,44$ bitų seka. Tačiau ne visiems tekstams Šenono rėžis yra pasiekiamas. Geriausio T kodo ilgis 23 bitai, pvz. $a \mapsto 0,\ b \mapsto 100,\ d \mapsto 110,\ k \mapsto 111,\ r \mapsto 101\,.$

Pasiūlymas: panagrinėkite tekstą $T' = abrakadabrakaada$.

2pvz. (Ligos rizikos faktoriai)

Atliekamas tyrimas, siekiant nustatyti kokie faktoriai įtakoja galimybę susirgti tam tikra liga. Ligos požymį žymėsime kintamuoju Y, įgyjančiu reikšmes S ir N (serga, neserga). Nagrinėjami trys faktoriai:

X_1 – paciento lytis, galimos reikšmės $\{M, V\}$;
X_2 – rūkymas, galimos reikšmės $\{taip, ne\}$;
X_3 – kraujospūdis, galimos reikšmės {*mažas,normalus,didelis*} .

Kuris faktorius labiausiai įtakoja polinkį susirgti? Norėdami atsakyti į šį klausimą, turime išsiaiškinti, kaip pasikeičia Y entropija, vieno ar kito požymio atžvilgiu. Kitaip sakant, turime palyginti tris tarpusavio informacijas

$$I(Y, X_i) = H(Y) - H(Y|X_i)\,, \quad i = 1, 2, 3.$$

100 pacientų tyrimo duomenys pateikti lentelėje.

2pvz. (Ligos rizikos faktoriai)

Paciento lytis (X_1)	*Rūkymas* (X_2)	*Kraujospūdis* (X_3)	*Ligos požymis* (Y)	Pacientų skaičius
M	*ne*	*mažas*	*N*	5
M	*ne*	*mažas*	*S*	2
M	*ne*	*normalus*	*N*	10
M	*ne*	*didelis*	*S*	6
M	*taip*	*mažas*	*N*	4
M	*taip*	*mažas*	*S*	2
M	*taip*	*normalus*	*N*	8
M	*taip*	*didelis*	*N*	1
M	*taip*	*didelis*	*S*	8
V	*ne*	*normalus*	*N*	8
V	*ne*	*didelis*	*S*	10
V	*ne*	*mažas*	*N*	2
V	*taip*	*mažas*	*S*	7
V	*taip*	*normalus*	*N*	6
V	*taip*	*normalus*	*S*	5
V	*taip*	*didelis*	*S*	16

2pvz. (Ligos rizikos faktoriai)

Atsitiktinio dydžio Y skirstinys yra

Y	N	S
P	$0,44$	$0,56$

Vadinasi, $H(Y) = h(0,44) \approx 0,989587521$
Skaičiuosime $H(Y|X_1)$.

$$\begin{aligned} & H(Y|X_1 = M) = \\ & P(Y = N|X_1 = M)\log_2 \frac{1}{P(Y = N|X_1 = M)} \\ + \quad & P(Y = S|X_1 = M)\log_2 \frac{1}{P(Y = S|X_1 = M)} \\ = \quad & h\left(\frac{28}{46}\right) \approx 0,965636133 \end{aligned}$$

Analogiškai $H(Y|X_1 = V) = h\left(\frac{16}{54}\right) \approx 0,876716289$

2pvz. (Ligos rizikos faktoriai)

Dabar randame $H(Y|X_1)$

$$\begin{aligned} & H(Y|X_1) = \\ & P(X_1 = M)H(Y|X_1 = M) + P(X_1 = V)H(Y|X_1 = V) \\ = \quad & 0,46 \cdot h\left(\frac{28}{46}\right) + 0,54 \cdot h\left(\frac{16}{54}\right) \approx 0,917619417 \end{aligned}$$

Analogiškai $H(Y|X_2) \approx 0,945171922$ ir $H(Y|X_3) \approx 0,499226424$.
Dabar jau galime rasti ieškomąsias tarpusavio informacijas

$$I(Y, X_1) \approx 0,071968104\,,$$

$$I(Y, X_2) \approx 0,044415599\,,$$

$$I(Y, X_3) \approx 0,490361097\,.$$

Absoliučiai tolydžiojo a. d. entropija

Tolydžiųjų atsitiktinių dydžių entropija apibrėžiama kitaip. Skiriasi ir jos interpretacija.

Apibrėžimas. *Tolydžiojo atsitiktinio dydžio X su tankio funkcija $p_X(x)$ entropija $H(X)$ yra lygi*

$$H(X) = -\int_{-\infty}^{\infty} p_X(x) \log_2 p_X(x)\, dx\,.$$

Iš karto reikia pažymėti, kad toldžiojo atsitiktinio dydžio entropijos negalima interpretuoti kaip vidutinės informacijos, nes šiuo atveju $P(X = x) = 0$, nepriklausomai nuo tankio funkcijos $p_X(x)$ reikšmės. Be to, toldžiojo atsitiktinio dydžio entropija gali būti ir neigiama.

Gauso a. d. entropija

Tegul X yra normalusis (kitaip: Gauso) atsitiktinis dydis su vidurkiu a ir dispersija $\sigma^2,\ \sigma > 0$. Tokio atsitiktinio dydžio tankio funkcija yra

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}.$$

Rasime jo entropiją. Pagal apibrėžimą

$$H(X) = \int_{-\infty}^{\infty} p_X(x)\left(\frac{(x-a)^2}{2\sigma^2}(\log_2 \mathrm{e}) + \log_2(\sqrt{2\pi}\sigma)\right) dx$$

$$= \frac{\log_2 \mathrm{e}}{2\sigma^2}\int_{-\infty}^{\infty}(x-a)^2 p_X(x)\,dx + \log_2(\sqrt{2\pi}\sigma)\int_{-\infty}^{\infty} p_X(x)\,dx$$

$$= \frac{\log_2 \mathrm{e}}{2\sigma^2}\cdot\sigma^2 + \log_2(\sqrt{2\pi}\sigma)\cdot 1 = \frac{1}{2}\log_2(2\mathrm{e}\pi\sigma^2).$$

Absoliučiai tolydaus a. d. entropijos įvertis

"Tolydžioji" Gibbs'o nelygybės versija yra

$$\int_{-\infty}^{\infty} p_X(x) \log_b \left(\frac{p_Y(x)}{p_X(x)} \right) \, dx \leq 0 \, .$$

Čia $b > 1$, o $p_X(x)$ ir $p_Y(x)$ - bet kokios tankio funkcijos. Pasinaudoję šia nelygybe, galime įrodyti tokį teiginį.

Teorema. *Jei absoliučiai tolydus atsitiktinis dydis* X *turi baigtinę dispersiją* $\mathbf{D}X = \sigma^2$, $\sigma > 0$, *tai jo entropija*

$$H(X) \leq \frac{1}{2} \log_2(2e\pi\sigma^2) \, .$$