INFORMACIJA I NJENA MERA opservacije i dogadjaji

- Šta ćemo naučiti u ovoj lekciji:
- Opservacija slučajne veličine ili dogadjaji vezani za jednu slučajnu veličinu predstavljaju informaciju
- Količina informacije dobijena na osnovu observacija slučajne veličine ili dogadjaja se meri na osnovu rezultujuće promene neodredjenosti
- Entropija i mera informacije su tesno povezani.

INFORMACIJA I NJENA MERA

- Šta je informacija i kako se ona meri?
- Osnovna ideja je da je ona vezana za promenu neodredjenosti.
- Mera informacije, stoga mora biti srazmerna količini promene neodredjenosti.
- Počnimo naša razmatranja od sistema mogućnosti (S,P), reprezentovanog slučajnom promenljivom X, koja uzima vrednosti $x \in S$, sa verovatnoćom $p_X(x)$. Ova slučajna promenljiva opisuje jedan eksperiment u kome su ishodi neodredjeni. Mera neodredjenosti u vezi sa ovim eksperimentom je data entropijom
- $$H(X) = -\sum_{x \in S} p_X(x) \log p_X(x). \qquad (1.15)$$

INFORMACIJA I NJENA MERA

Kada se izvrši eksperiment, mi observiramo odredjenu vrednost $x \in S$. Nakon toga više nema neodredjenosti.

Prethodna neodredjenost H(X) redukovana na aposteriornu neodredjenost 0. Razlika H(X) – 0 = H(X) jednaka je količini informacije dobijene izvodjenjem ovog eksperimenta.

Stoga je entropija jedne slučajne veličine jednaka količini informacije dobijene observiranjem jedne konkretne njene realizacije.

Ova ideja nam nameće dva važna zapažanja:

- Budući da je informacija jednaka promeni entropije, ona se meri u istim jedinicama kao i entropija, dakle ako je 2 osnova logaritma, jedinica je jedan bit.
- Dobijena količina informacije je ista za sve moguće realizacije slučajne veličine X, bez obzira da li je verovatnoća neke realizacije velika ili mala.

INFORMACIJA I NJENA MERA

- Primer 1.15 *Binarna slučajna promenljiva*
- Neka je zadata binarna slučajna veličina X, koja prima vrednosti 0 i 1 sa verovatnoćama p i q=1-p, respektivno.
- Opserviranjem jednog ishoda binarnog eksperimenta rezultuje u dobijanju količine informacije jednake
- $H(X)= - p \log p – q \log q$.
- Npr. u fer bacanju novčića, opserviranjem ishoda bacanja se dobija količina informacije od 1 bita.

INFORMACIJA I NJENA MERA

- Generališimo sada ovo razmatranje. I dalje ćemo posmatrati slučajnu promenljivu X vezanu za sistem mogućnosti (S,P).
- Količina neodredjenosti je i dalje H(X). Sada ćemo obaviti eksperiment samo parcijalno. Nećemo opservirati direktno slučajnu veličinu X, već samo neki dogadjaj $E \subseteq S$. I ovo predstavlja neku informaciju. Šta je sada njena mera?
- Opservacija dogadjaja E menja slučajnu veličinu X u uslovnu promenljivu $X|E$, koja se odnosi na novi verovatnosni sistem mogućnosti (E, P_E). Ovde P_E označava uslovne verovatnoće

$$p_{X|E}(x) = \frac{p_X(x)}{p_X(E)}, \text{ za svako } x \in E.$$

INFORMACIJA I NJENA MERA

- Ova nova situacija, nastala opservacijom dogadjaja E, ima neodredjenost jednaku uslovnoj entropiji

$$H(X|E) = -\sum_{x\in E} p_{X|E}(x)\log p_{X|E}(x).$$

- Prema tome opserviranje dogadjaja E, menja neodredjenost sa početne H(X) u H($X|E$).
- Količina informacije dobijena tom prilikom je jednaka

H(X) - H($X|E$).

- Videćemo da ovo ne mora biti u opštem slučaju dobitak, budući da H($X|E$) može biti i veća od H(X), tako da opserviranje dogadjaja E može da poveća neodredjenost, što vodi definiciji negativne informacije.

INFORMACIJA I NJENA MERA

Kada se E svodi na opservaciju tačne vrednosti jedne od mogućnosti X, tada je H($X|E$) = 0, pa je

$$H(X) - H(X|E) = H(X) – 0 = H(X)$$

količina informacije dobijena tom prilikom jednaka entropiji.

Zato se često entropija neke slučajne veličine naziva i informacija o samoj sebi (selinformation). Odnosno entropija jedne slučajne veličine je u srednjem jednaka količini informacije potrebne za opis te slučajne veličine.

Da bi smo nastavili sa razvojem koncepta informacije nužno je da uvedemo dva nova pojma:

- relativnu entropiju i uzajamnu informaciju

UZAJAMNA INFORMACIJA

Posmatrajmo združeni verovatnosni sistem mogućnosti$(S_1 \times$

UZAJAMNA INFORMACIJA

- Medjutim opservacija Y=y, menja i raspodelu verovatnoće slučajne veličine X, tako što ona sada postaje uslovna raspodela

$$p_{X|y}(x,y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}, \text{za svako } x \in S_1.$$

- Ovo znači da se takodje menja i neodredjenost u pogledu slučajne promenljive X. Stoga opservacija Y=y, sadrži informaciju u odnosu na X. Količina ove informacije iznosi

$$i(y;X) = H(X) - H(X|y)$$

- Ovu veličinu ćemo zvati lokalna uzajamna informacija slučajne veličine Y i X, a govori o tome koliko jedna opservacija Y=y daje informacije o drugoj slučajnoj veličini X.

UZAJAMNA INFORMACIJA

- Predjimo sada sa pojedinačnih opservacija na očekivane vrednosti lokalnih uzajamnih informacija

$$I(X;Y) = \sum_y p_Y(y) i(y;X) = \sum_y p_Y(y)(H(X) - H(X|y)) = H(X) - H(X|Y).$$

- Veličina I(X;Y) se naziva uzajamna informacija izmedju slučajnih veličina X i Y. Ona predstavlja jedan od centralnih pojmova teorije informacija. Striktno govoreći ona nije informacija već očekivana vrednost količine informacije koju dobijamo o X kada opserviramo vrednosti Y.
- Na osnovu Korolarija 1.3 koji tvrdi da je uvek

$$H(X|Y) \leq H(X)$$

UZAJAMNA INFORMACIJA

TEOREMA 1.8

$$I(X;Y) \geq 0, \qquad (1.29)$$

dok je $I(X;Y) = 0$, ako i samo ako us X I Y medjusobno nezavisne slučajne veličine.

- Podsetimo se da lokalna uzajamna informacija $i(y;X)$ može biti i negativna, dok je uzajamna informa kao njeno matematičko očekivanje uvek pozitivna.
- Na osnovu svojstva H(X|Y)=H(X,Y) - H(Y), dobijamo da važi

$$I(X;Y) = \mathrm{H(X)} + \mathrm{H(Y)} - \mathrm{H(X,Y)}. \qquad (1.29a)$$

INFORMACIJA I NJENA MERA

Iz simetrije izraza (1.29a) sledi sledeća

TEOREMA 1.9

$$I(X;Y) = I(Y;X).$$

Sumirajmo sada najznačajnija svojstva uzajamne informacije i njen odnos prema entropiji

TEOREMA 1.9a

$$I(X;Y) = \mathrm{H(X)} - \mathrm{H(X|Y)}$$
$$I(X;Y) = \mathrm{H(Y)} - \mathrm{H}(Y|\mathrm{X})$$
$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$
$$I(X;Y) = I(Y;X)$$
$$I(X;X) = H(X)$$

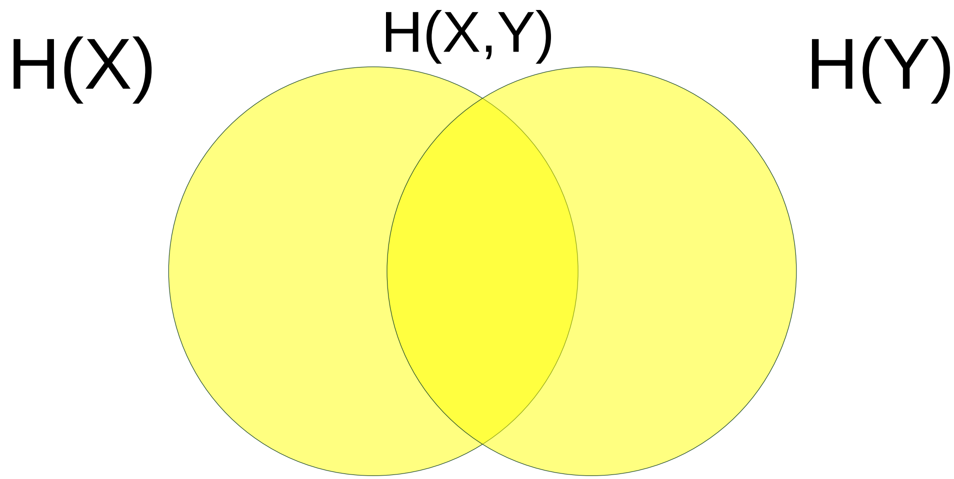
Grafička prezentacija informaciono-teorijskih veličina

- Raymond W, Yueng , “A new Outlook on Shannon Information Measures”, *IEEE Transactions on IT, Vol.37, No.3, May 1991.*

- X U Y = H(X,Y)
- X = H(X)
- Y = H(Y)
- $X \cap Y = I(X;Y)$
- X – Y = H(X l Y) $X - Y = X \cap Y^c$
- Y – X = H(Y l X)
- $(X \cap Y)^c = H(X \mid Y) + H(Y \mid X)$

, → U
; → ∩
l → -

Grafička prezentacija informaciono-teorijskih veličina

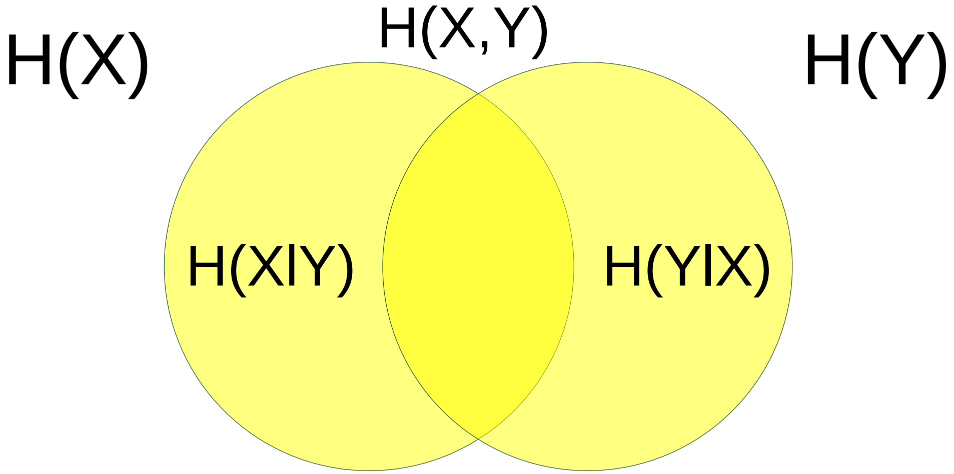


Dve slučajne varijable

$H(X,Y) \leq H(X)+H(Y)$

$0 \leq H(X) \leq \log_2 n$ n – broj mogućih vrednosti X

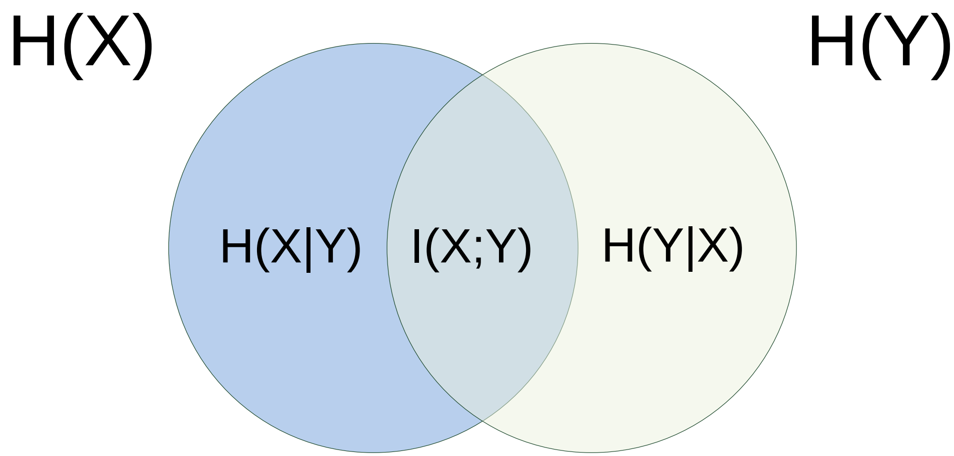
Grafička prezentacija informaciono-teorijskih veličina



Dve slučajne varijable

$H(X|Y) \leq H(X)$ $H(Y|X) \leq H(Y)$

Grafička prezentacija informaciono-teorijskih veličina

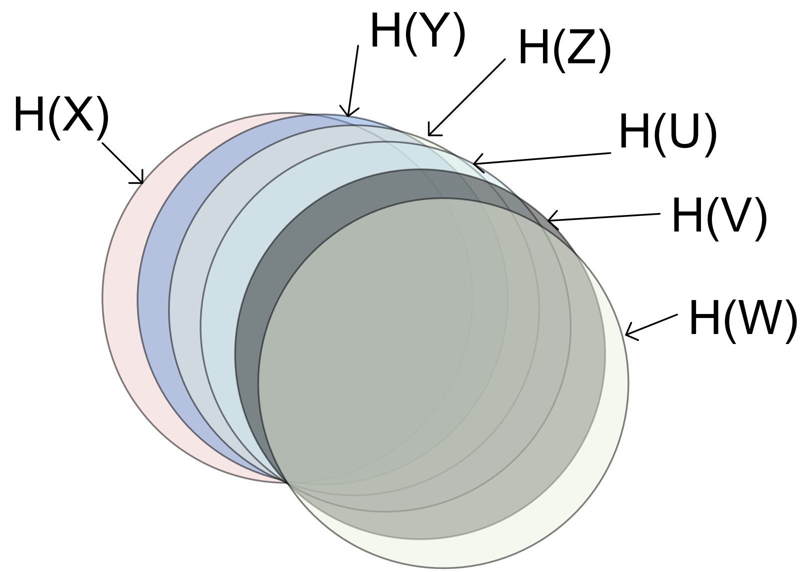


Dve slučajne varijable

$I(X;Y)=H(X)-H(X|Y)$ $I(X;Y)=H(Y)-H(Y|X)$

Uzajamna informacija

Grafička prezentacija informaciono-teorijskih veličina



Markovljev lanac X→Y→Z→U→V→W

RELATIVNA ENTROPIJA

- Posmatrajmo dva verovatnosna sistema mogućnosti nad istim skupom mogućnosti S. Ako su X i Y korespodentne slučajne promenljive, tada se veličina

$$K(P_X, P_Y) = \sum_{x \in S} p_X(x) \log \frac{p_X(x)}{p_Y(x)} \qquad (1.31)$$

- naziva relativna entropija ili Kulbak Lejblerova divergencija (Kulbak-Lejblerovo rastojanje) izmedju raspodele dve slučajne veličine X i Y.
- Iako se u jednom od naziva koristi rastojanje, primetimo da ne važi uslov simetričnosti, naime

$$d(P_X, P_Y) \neq d(P_Y, P_X).$$

RELATIVNA ENTROPIJA

Divergencija je jednaka nuli

$$K(P_X, P_Y) = K(P_Y, P_X) = 0$$

ako i samo ako su

$$P_X = P_Y.$$

Pošto se usvaja konvencija da je 0 log(0/0)=0,

0 log(0/q)=0, p log (p/0)= ∞,

K(P_X, P_Y)=∞, ukoliko postoji x za koje je $p_X(x) > 0$, a $p_Y(x)$=0.

RELATIVNA ENTROPIJA

- Posmatrajmo združeni verovatnosni sistem mogućnosti $(S \times S, P)$ kome je asociran par slučajnih veličina $X, Y \in S$.
- Marginalni sistemi (S, P_X) i (S, P_Y) su indukovani marginalnim raspodelama verovatnoća

$$p_X(x) = \sum_y p_{X,Y}(x,y), \quad p_Y(y) = \sum_x p_{X,Y}(x,y).$$

- Označimo sa $P_X \cdot P_Y$ raspodelu verovatnoća $p_X(x) \cdot p_Y(y)$.
- Tada važi

$$I(X;Y) = I(Y;X) = K(P, P_X \cdot P_Y).$$

RELATIVNA ENTROPIJA

$$I(X;Y) = \sum_y p_Y(y) i(y;X) = \sum_y p_Y(y)(H(X) - H(X|y)) = H(X) - H(X|Y).$$

Ako u gornji izraz zamenimo vrednosti za H(X) i H(X|Y) dobijamo

$$\begin{aligned}
I(X;Y) &= -\sum_x p_X(x)\log p_X(x) + \sum_{x,y} p_Y(y) p_{X|Y}(x,y)\log p_{X|Y}(x,y) \\
&= \sum_{x,y} p_{X,Y}(x,y)\left(\log \frac{p_{X,Y}(x,y)}{p_Y(y)} - \log p_X(x)\right) \\
&= \sum_{x,y} p_{X,Y}(x,y)\log \frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)}.
\end{aligned}$$

RELATIVNA ENTROPIJA

- Na osnovu ovih uvida nam je jasno da uzajamna informacija I(X;Y) meri stepen divergencije združene raspodele verovatnoće para (X,Y) od raspodele verovatnoće ovog para u slučaju njihove nezavisnosti.
- Stoga se često sa opravdanjem uzajamna informacija uzima kao mera medjuzavisnosti dve slučajne veličine koje uzimaju vrednosti nad istim skupom.