CANAIS DE COMUNICAÇÃO E TEORIA DE INFORMAÇÃO

Ramon e Leslie

UNIVERSIDADE ESTADUAL DE CAMPINAS

lkbc456@gmail.com

10 de junho de 2018

Sumário

Mensagens discretas e entropia

Para a transmissão em canais com erros precisamos de um modelo das mensagens que queremos transmitir. Esses modelos são geralmente divididos em strings de mensagens, normalmente usados na forma binária.

Definições iniciais

Em receptores discretos sem memória, a saída é uma sequência com as mesmas propriedades do seu modelo. Cada saída, X, tem valores finitos no alfabeto do modelo $\{x_1, x_2, ..., x_n\}$. Já a probabilidade de x_j acontecer é $P\left(x_j\right) = p_j$, com a distribuição de probabilidade sendo $Q\left(x\right) = \{p_1, p_2, ..., p_n\}$.

ENTROPIA

Def1:ENTROPIA

A entropia é a medida da incerteza de uma variável aleatória e para um receptor discreto sem memória, X, é definida como:

$$H(X) = -\sum_j p_j.logp_j = E[-logP(X)]$$

Normalmente os logaritmos têm base 2 e H é representada em bits.

ENTROPIA

Def1:ENTROPIA

A entropia é a medida da incerteza de uma variável aleatória e para um receptor discreto sem memória, X, é definida como:

$$H(X) = -\sum_j p_j.logp_j = E[-logP(X)]$$

Normalmente os logaritmos têm base 2 e H é representada em bits.

Expectativa(E)

O valor esperado da variável g(X):

$$E[g(X)] = \sum_{x \epsilon \lambda} g(X)p(X)$$

Exemplo: Entropia

Se: $X = \begin{cases} 1 \text{ com probabilidade p} \\ 0 \text{ com probabilidade 1-p} \end{cases}$

$H(X) = -plog_2 p - (1-p)log_2(1-p)$
$E(X) = 1.p + 0.(1-p)$

ENTROPIA

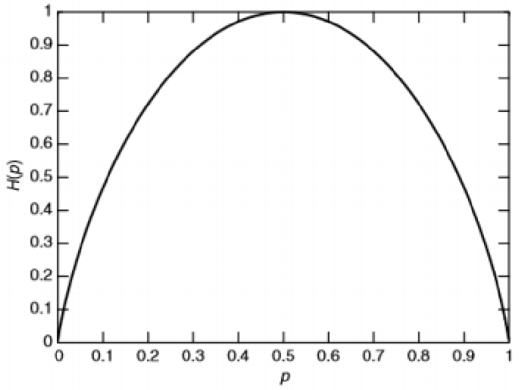

Figura: $H(X)xp$ - Função da entropia binária

Lemma[CT]2.1.1

$$H(X) \geq 0$$

PROVA: $0 \leq p(x) \leq 1 \Rightarrow \log(\frac{1}{p(x)}) \geq 0$

Lemma[CT]2.1.1

$$H(X) \geq 0$$

PROVA: $0 \leq p(x) \leq 1 \Rightarrow \log(\frac{1}{p(x)}) \geq 0$

Lema[CT]2.1.2 Mudança de base

$$H_b = (log_b a)H_a(X)$$

ENTROPIA CONJUNTA E ENTROPIA CONDICIONAL

Estendendo a definição de entropia de uma única variável para um par de variáveis (X, Y), quando as consideramos com um único vetor de variáveis aleatórias:

ENTROPIA CONJUNTA E ENTROPIA CONDICIONAL

Estendendo a definição de entropia de uma única variável para um par de variáveis (X, Y), quando as consideramos com um único vetor de variáveis aleatórias:

Def1:Entropia Conjunta $H(X, Y)$

A entropia conjunta $H(X, Y)$ de um par com distribuição conjunta de probabilidade $p(x, y)$ é definida como:

$$H(X, Y) = \sum_{x \epsilon X} \sum_{y \epsilon Y} p(x, y) logp(x, y) = -E[logp(x, y)]$$

ENTROPIA CONJUNTA E ENTROPIA CONDICIONAL

Def2:Entropia Condicional $H(Y|X)$

Se (X, Y) $p(x, y)$, a entropia condicional H(Y|X) é definida como:

$$H(X|Y) = \sum_{x \epsilon X} p(x, y) logp(x) H(Y|X = x) = -E[log(Y|X)]$$

ENTROPIA CONJUNTA E ENTROPIA CONDICIONAL

Def2:Entropia Condicional $H(Y|X)$

Se (X, Y) $p(x, y)$, a entropia condicional H(Y|X) é definida como:

$$H(X|Y) = \sum_{x \epsilon X} p(x, y) logp(x) H(Y|X = x) = -E[log(Y|X)]$$

Teorema[CT]:Regra da Cadeia

$$H(X, Y) = H(X) + H(Y|X)$$

Exemplo: Entropia conjunta e condicional.

Se (X, Y) tem a seguinte distribuição conjunta.

$Y \backslash X$	1	2	3	4
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
4	$\frac{1}{4}$	0	0	0

$H(Y|X) = \frac{13}{8}$ e $H(X, Y) = \frac{27}{8}$
Daí: $H(X) = \frac{14}{8}$

Um canal de informação é um modelo de comunicação ou sistema relacionado em que a entrada é uma mensagem e a saída é uma reprodução imperfeita da entrada. O canal pode ter várias restrições nas mensagens e problemas de transmissão.

Informação mútua e capacidade de canais discretos

Em um canal discreto sem memória a entrada e saída são sequências de símbolos, cuja saída depende somente da entrada atual.E conveniente representar o canal pela matriz de transição $Q(Y|X) = [pji]$.

$$Q(Y) = Q(X)Q(Y|X)$$

Em um canal discreto sem memória a entrada e saída são sequências de símbolos, cuja saída depende somente da entrada atual.E conveniente representar o canal pela matriz de transição $Q(Y|X) = [pji]$.

$$Q(Y) = Q(X)Q(Y|X)$$

Def1:Informação mútua $I(X, Y)$

Definida como a quantidade de informação de X dada por Y

$$I(X, Y) = E[\frac{logP(x|y)}{P(y)}] = \sum_j p(x_j) \sum_i p_{ji}[logp_{ji} - logP(y_i)] =$$

$$\sum_j \sum_i p(x_j, y_i) log(\frac{p(x_j, y_i)}{p(x_j)p(y_i)})$$

RELAÇÃO ENTROPIA E INFORMAÇÃO MÚTUA

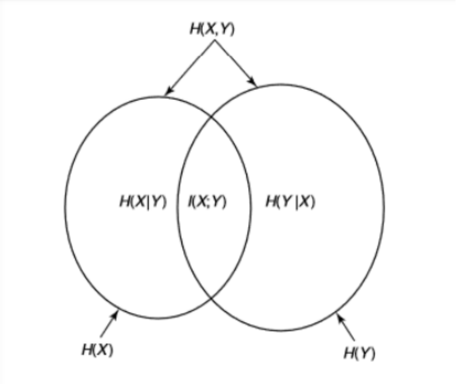

Figura:

$$I(X,Y) = H(Y) - H(Y|X) = H(X) - H(X|Y) = H(X) + H(Y) - H(X,Y)$$

Exemplo: Entropia e Informação Mutua
Se (X, Y) tem a seguinte distribuição conjunta.

$Y \backslash X$	1	2	3	4
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
4	$\frac{1}{4}$	0	0	0

$H(X|Y) = \frac{11}{8}$ e $H(X) = \frac{14}{8}$
Daí: $I(X;Y) = \frac{3}{8}$

Def: Capacidade de um canal discreto

Já a capacidade de um canal discreto, $C(Y|X)$ é o máximo de I com respeito a P(X).
Em vários casos interessantes, a simetria de transições entre X e Y indicam que a capacidade máxima é obtida com entradas simétricas.

Def: Capacidade de um canal discreto

Já a capacidade de um canal discreto, $C(Y|X)$ é o máximo de I com respeito a P(X).
Em vários casos interessantes, a simetria de transições entre X e Y indicam que a capacidade máxima é obtida com entradas simétricas.

BSC:Canal binário simétrico

Para canais em que erros ocorrem com uma probabilidade p em dados binários. Sua matriz de transição é

$$Q = \begin{vmatrix} 1-p & p \\ p & 1-p \end{vmatrix}$$

para p=0.5 a capacidade é 0

BEC

Canal que apaga símbolos se caso eles tenham uma saída com um carácter que não é o que foi mandado. Matriz de transição:

$$\begin{bmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{bmatrix}$$

em que $C = 1 - p$

Códigos e capacidade de canal

A importância da capacidade é relacionada a teorias de códigos, que indicam que k bits de mensagens podem se comunicar usando o canal um pouco mais do que $n = \frac{k}{C}$ vezes.

Então, um codificador mapeia k bits de mensagem com n símbolos codificados, usando um código consistindo de 2^k vetores

Distribuição binomial

$$A(z) = 1 + \sum_{w>0} 2^{k-n} \binom{n}{w} z^w$$

Códigos e capacidade de canal

A importância da capacidade é relacionada a teorias de códigos, que indicam que k bits de mensagens podem se comunicar usando o canal um pouco mais do que $n = \frac{k}{C}$ vezes.
Então, um codificador mapeia k bits de mensagem com n símbolos codificados, usando um código consistindo de 2^k vetores

Distribuição binomial

$$A(z) = 1 + \sum_{w>0} 2^{k-n} \binom{n}{w} z^w$$

LEMA[*JH*]4.2.2

$$\sum_{j=0}^{m} \binom{n}{j} \leq 2^{nH(\frac{m}{n})}$$

Probabilidade de erro

Quando $R < C$, existe uma constante positiva $E(R)$

$$P_{err} < 2^{-NE(R)}$$

Sendo que R é a taxa de bits por simbolo do canal, $N = \frac{(Z+1)w}{Z}$ e $Z = \sqrt{4p(1-p)}$.

Teorema 1

Para taxas $R < R_0$ e quaisquer blocos de tamanho n, existem blocos de códigos tal que a probabilidade de erro em uma BSC satisfaz:

$$P_{err} < 2^{-n(R_0 - R)}$$

onde

$$R_0 = 1 - log(1 + Z)$$

Relação entre entropia relativa e informação mutua

Entropia relativa é a medida da distância entre duas distribuições. Logo,é a medida da ineficiência em assumir que a distribuição é q quando na verdade è p.

Def:1

Entropia relativa ou distância de Kullback-Leibler entre $p(x)$ e $q(x)$.

$$d(p||q) = \sum_{x \epsilon X} p(x) log \frac{p(x)}{q(x)} = E_p log \frac{p(x)}{q(x)}$$

Relação entre entropia relativa e informação mutua

Entropia relativa é a medida da distância entre duas distribuições. Logo,é a medida da ineficiência em assumir que a distribuição é q quando na verdade è p.

Def:1

Entropia relativa ou distância de Kullback-Leibler entre $p(x)$ e $q(x)$.

$$d(p||q) = \sum_{x \epsilon X} p(x) log \frac{p(x)}{q(x)} = E_p log \frac{p(x)}{q(x)}$$

Relação com informação mútua:

$$I(x, y) = D(p(x, y)||p(x)p(y))$$

Regras de cadeia para entropia, entropia relativa e informação mútua

Def 1

A informação mútua condicional de variáveis aleatórias X e Y dado z é definida como:

$$I(X;Y|Z) = H(X|Z) - H(X|Y;Z) = E_{p(x,y,z)} log \frac{p(X;Y|Z)}{p(X|Z)p(Y|Z)}$$

Teorema2.6.2[CT]

Desigualdade de Jensen Se f é convexa e X é uma variável aleatória,

$$Ef(X) \geq f(EX)$$

Se f é estritamente convexo e existe a igualdade na relação, $X = EX$, X é constante.

Teorema2.6.2[CT]

Desigualdade de Jensen Se f é convexa e X é uma variável aleatória,

$$Ef(X) \geq f(EX)$$

Se f é estritamente convexo e existe a igualdade na relação, $X = EX$, X é constante.

Teorema 2.6.3[CT]

Desigualdade de informação:Se $p(x)$, $q(x)$, $x\lambda X$, forem as funções de probabilidade. Então:

$$D(p||q) \geq 0$$

, existindo igualdade se $p(x) = q(x)$.

Corolários

- $I(X;Y) \geq 0$, com igualdade somente se X e Y são independentes.
- $D((p(y|x))||q(y|x)) \geq 0$, com igualdade somente para $p(y|x) = q(y|x)$, onde $p(x) > 0$.
- $I(X;Y|Z) \geq 0$, com igualdade somente se X e Y são condicionalmente independente dado Z.

Desigualdade de Jensen e suas consequências

Teorema2.6.4[CT]

$H(X) \leq log|X|$, onde $|X|$ denota o número de elementos de X, com a igualdade só existindo se X tiver uma distribuição uniforme.

Desigualdade de Jensen e suas consequências

Teorema2.6.4[CT]

$H(X) \leq log|X|$, onde $|X|$ denota o número de elementos de X, com a igualdade só existindo se X tiver uma distribuição uniforme.

Teorema 2.6.5[CT]

Condicionamento reduz entropia

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

, com igualdade somente se X e Y independentes.

Teorema2.6.6$[CT]$

Limitante de independência da entropia. Se,$(X_1, X_2, ..., X_n)$ pode ser definido por $p(x_1, x_2, ..., x_n)$:

$$H(X_1, X_2, ..., X_n) \leq \sum_{i=1}^{n} H(X_i)$$

Final