

Teoria da Informação

Ney Lemke

lemke@exatas.unisinos.br

Ciência da Computação
Mestrado em Computação Aplicada
Unisinos

Plano do Curso

- Revisão Conceitos Matemáticos
- Medidas de Incerteza
- Canais de Comunicação
- Teorema de Shannon
- Técnicas de Compressão de Dados
- Criptografia

Revisão Matemática

Teoria dos Conjuntos

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$
- $A - B = A \cap \overline{B}$

Espaço de Probabilidades

Definição Um espaço de probabilidade é um par (S, P) no qual S é um conjunto não vazio e $P : S \rightarrow [0, 1]$ é uma função satisfazendo

$$\sum_{u \in S} P(u) = 1$$

S é um conjunto de possibilidades.

P é uma função que associa a cada possibilidade um probabilidade.

Partição

Uma partição de S , S' é um conjunto de subconjuntos de S designados por P_1, \dots, P_n tal que:

- $P_i \cap P_j = \emptyset$
- $\bigcup_{i=1}^m P_i = S$

Eventos

Seja (S, P) um espaço de probabilidade finito. Um evento neste espaço é um subconjunto de S . A probabilidade de ocorrência de E é

$$P(E) = \sum_{u \in E} P(u)$$

Eventos Mutuamente Exclusivos

Eventos E_1 e E_2 em um espaço de probabilidade (S, P) são mutuamente exclusivos se

$$P(E_1 \cap E_2) = 0$$

Resultados

- Se $E_1 \subset E_2$ $P(E_1) \leq P(E_2)$
- Se $E_1 \cap E_2 = \emptyset$, $F_1 \subset E_1$, $F_2 \subset E_2$ então $F_1 \cap F_2 = \emptyset$
- Considere E_1, \dots, E_m eventos mutuamente exclusivos:

$$P\left(\bigcup_{i=1}^m E_i\right) = \sum_{i=1}^m P(E_i)$$

- Se $E \subset F \subset S$ então $P(F - E) = P(F) - P(E)$
- $P(E \cup F) + P(E \cap F) = P(E) + P(F)$

Exemplo

Suponha que em uma população 40% das pessoas tem cabelo ruivo, 25% é tuberculoso e 15% possui ambos. Qual é a percentagem que não tem nenhuma das características?

Exercício

Em uma dada população 25% das pessoas é pequena e suja, 35% das pessoas é grande e limpa 60% são pequenos. Que percentagem é suja?

Probabilidade Condicional

Definição Suponha (S, P) é um espaço de probabilidade finito. $E_1, E_2 \subset S$ e $P(E_2) \neq 0$. A probabilidade condicional de E_1 dado E_2 é:

$$P(E_1|E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)}$$

Exemplos

Uma urna contém 5 bolas vermelhas, 12 verdes e 8 amarelas. Três delas são escolhidas sem troca.

- Qual é a probabilidade que uma bola vermelha, verde e amarela sejam escolhidas ?
- Qual é a probabilidade que a última bola escolhida seja verde ?

Exercício

Uma urna contém 6 bolas vermelhas, 5 bolas verdes e 3 bolas amarelas. Duas são escolhidas sem troca. Qual é a probabilidade que pelo menos uma seja amarela?
Considere a situação em que três bolas são tomadas.

Eventos Independentes

Definição Seja (S, P) um espaço de probabilidades finito e $E_1, E_2 \subset S$. Os eventos E_1 e E_2 são independentes se e somente se:

$$P(E_1 \cap E_2) = P(E_1)P(E_2)$$

Nota: Se E_1 e E_2 são independentes $P(E_1|E_2) = P(E_1)$.

Exercício

Admita que E e F sejam eventos independentes em um espaço de probabilidade finito (S, P) . Mostre que:

- E e $S - F$ são independentes
- $S - E$ e $S - F$ são independentes.

Tentativas de Bernoulli

Definição Suponha $n \geq 0$ e k inteiros. $C_{n,k}$ o coeficiente binomial é definido por:

$$C_{n,k} = \frac{n!}{k!(n-k)!}$$

Alfabeto

Definição: Um alfabeto A é um conjunto não vazio e finito. Uma palavra de comprimento n sobre um alfabeto A é uma sequência de tamanho n formada por elementos de A .

Exemplos Sejam α e β símbolos distintos. Então: $C_{n,k} = |\{w, w \text{ é uma palavra de comprimento } n \text{ onde } \alpha \text{ aparece exatamente } k \text{ vezes em } w\}|$

Resultados

- $(x + y)^n = \sum_{k=0}^n C_{n,k} x^k y^{n-k}$

- $2^n = \sum_{k=0}^n C_{n,k}$

Tentativas de Bernoulli

Uma tentativa de Bernoulli é um experimento com dois possíveis resultados. Uma seqüência de tentativas de Bernoulli é um experimento com vários estágios, onde cada experimento é de Bernoulli com estágios independentes.

Teorema

Suponha que a probabilidade de sucesso em uma tentativa de Bernoulli é p . Então a probabilidade de k sucessos em uma sequência de n tentativas é:

$$C_{n,k} p^k (1 - p)^{n-k}$$

Exemplo

Uma urna contém 7 bolas vermelhas e 10 bolas azuis e 20 bolas são escolhidas com reposição. Qual é a probabilidade de:

- obter 4 bolas vermelhas
- obter pelo menos 4 bolas sejam vermelhas

Variáveis Aleatórias

Suponha que (S, P) é um espaço de probabilidade finito. Uma variável aleatória neste espaço é uma função de S nos números reais \mathbb{R} . Se $X : S \rightarrow \mathbb{R}$ é uma variável aleatória em (S, P) o valor esperado de X é:

$$E(X) = \sum_{u \in S} X(u)P(u)$$

Exemplo

Um experimento consiste de n tentativas de Bernoulli com probabilidade p de sucesso em cada tentativa. Seja “ X =número de sucessos”. Considere $n = 3$ $p = 1/2$ e determine $E(X)$.

Teorema

$$E \left(\sum_{k=1}^n a_k x_k \right) = \sum_{k=1}^n a_k E(x_k)$$

Teorema

O valor esperado para o número de sucessos em n tentativas de Bernoulli com probabilidade p é:

$$\sum_{k=0}^n k C_{n,k} p^k (1-p)^{(n-k)} = np$$

Lei dos Grandes Números

Teorema: Suponha que para uma tentativa de Bernoulli a probabilidade de sucesso é p . Para uma seqüência de n tentativas, considere X_n como sendo o número de sucessos em n tentativas. Suponha $\epsilon > 0$, então:

$$P \left(\left| \frac{X_n}{n} - p \right| < \epsilon \right) \rightarrow 1 \quad \text{quando } n \rightarrow \infty$$

Exemplo $p = 1/2$

- $P \left(\left| \frac{X_{10}}{10} - \frac{1}{2} \right| < \frac{1}{10} \right) = 0.24$
- $P \left(\left| \frac{X_{1000}}{1000} - \frac{1}{2} \right| < \frac{1}{1000} \right) = 0.95$

Entropia e Informação

Sistemas de Eventos

Definição Suponha que (S, P) é um espaço de probabilidades finito. Um sistema de eventos em (S, P) é uma coleção finita indexada $\mathcal{E} = [E_i, i \in I]$ de eventos mutuamente excludentes satisfazendo $P\left(\bigcup_{i \in I} E_i\right) = 1$.

- Como $E_i \cap E_j = \emptyset$ se $i \neq j$:

$$P\left(\bigcup_{i \in I} E_i\right) = \sum_{i \in I} P(E_i)$$

- Qualquer partição S é um sistema de eventos.

Exemplo

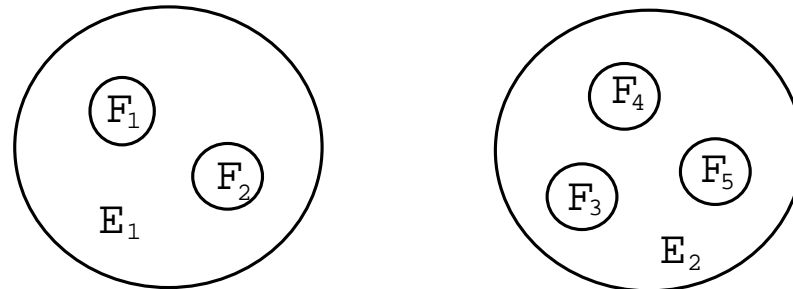
Considere as tentativas de Bernoulli $S = \{S, F\}^n$ se tomamos $E_k =$ exatamente k sucessos. Os E_k particionam S .

Observações

- É possível existir um sistema de eventos que não particione S , basta que existam eventos com probabilidades nulas.
- Seja \mathcal{F} um sistema de eventos de S . Podemos definir um novo sistema de eventos de probabilidade finito $S' = (\mathcal{F}, P)$.

Amalgamação

Suponha $\mathcal{E} = [E_i, i \in I]$ e $\mathcal{F} = [F_j, j \in J]$ são sistemas de eventos em um espaço de probabilidade finito. Nós dizemos que \mathcal{E} é uma amalgamação de \mathcal{F} se para todo $j \in J$ existir $i \in I$ tal que $P(E_i \cap F_j) = P(F_j)$.



Agrupamento

Suponha que $\mathcal{E} = [E_i, i \in I]$ e $\mathcal{F} = [F_j, j \in J]$ são sistemas de eventos em espaços de probabilidade finita (S, P) . O sistema agrupado de E com F denotado por

$$E \wedge F = [E_i \cap F_j; (i, j) \in I \times J].$$

Lema

Se (S, P) é um EPF, $E, F \subset S$ e $P(F) = 1$ então $P(E) = P(F \cap E)$.

Corolário

Se $\mathcal{F} = [F_j, j \in J]$ é um sistema de eventos em (S, P) e $E \subset S$ então

$$P(E) = \sum_{j \in J} P(E \cap F_j)$$

Teorema

Se \mathcal{E} e \mathcal{F} são sistemas de eventos em (S, P) então $\mathcal{E} \wedge \mathcal{F}$ é um sistema de eventos.

Eventos estatisticamente independentes

Suponha que $\mathcal{E} = [E_i, i \in I]$ e $\mathcal{F} = [F_j, j \in J]$ são sistemas de eventos em algum EPF (S, P) . \mathcal{E} e \mathcal{F} são estatisticamente independentes se e somente se E_i e F_j são eventos independentes para todo $i \in I$ e $j \in J$.

Exercício

Você possui dois dados um verde e um outro vermelho.

- $S_1 = \{ i \text{ “aparece no vermelho, } j \text{ no verde” } \}$
- $S_2 = \{ i \text{ “aparece em um dado, } j \text{ no outro” } \}$
- $S_3 = \{ i \text{ “a soma dos dados é } k \text{” } \}$
- $S_4 = \{ i \text{ “números pares nos dois lados”, “par no vermelho, ímpar no verde”, “ímpar no verde, par no vermelho”, “números ímpares nos dois” } \}$

Qual par de conjuntos não é uma amalgamação do outro?

Informação

Seja (S, P) um EPF se $E \subset S$ a auto informação de E é:

$$I(E) = -\log P(E) = \log \frac{1}{P(E)}$$

Notas:

- A base do logaritmo não está especificada, se a base é 2 a informação é medida em bits, se a base for e os logaritmos são medidos em “gnats”.
- Se $P(E) = 0$, $I(E) = \infty$

Informação Mútua

Suponha que $E, F \subset S$ e $\mathcal{E} = [E_i, i \in I]$ e $\mathcal{F} = [F_j, j \in J]$ são sistemas de eventos em (S, P) . A informação mútua entre E e F é

$$I(E, F) = \log \frac{P(E \cap F)}{P(E)P(F)}.$$

A informação mútua entre \mathcal{E} e \mathcal{F} é

$$I(\mathcal{E}, \mathcal{F}) = \sum_{i \in I} \sum_{j \in J} P(E_i \cap F_j) I(E_i, F_j)$$

A informação condicional entre E e F é

$$I(E|F) = -\log P(E|F) = -\log \frac{P(E \cap F)}{P(F)}.$$

Notas

- $I(E, F)=0$ se E e F são eventos independentes.
- Considere $E \wedge F$ como um conjunto novo de experimentos originário do conjunto S . $I(\mathcal{E}, \mathcal{F})$ é o valor médio da variável $I(E_i, F_j)$.
- Convenções:

$$0 \log \infty = 0 \log \frac{0}{0} = 0$$

Lema

$\ln x \leq x - 1$ para todo $x > 0$ com igualdade se e somente se $x=1$.

Teorema

Suponha que \mathcal{E} e \mathcal{F} são sistemas de eventos em um espaço de eventos finito (S, P) . $I(\mathcal{E}, \mathcal{F}) \geq 0$ com igualdade se e somente se \mathcal{E} e \mathcal{F} são estatisticamente independentes.

Exercícios

Suponha que E é um evento e \mathcal{E} um sistema de eventos em algum espaço de probabilidades.

- Mostre que se $P(E) > 0$ então $I(E, E) = I(E)$
- Mostre que $I(E, E) = 0$ se $P(E) > 0$.
- Em que condições $I(\mathcal{E}, \mathcal{E}) = 0$?

Assuma que E e F são eventos em um espaço de eventos em algum espaço de probabilidade, e $P(E), P(F) > 0$.

Mostre que:

$$I(E, F) = I(E) - I(E|F)$$

Entropia

Suponha que $\mathcal{E} = [E_i, i \in I]$ e $\mathcal{F} = [F_j, j \in J]$ são sistemas de eventos em algum EPF. A entropia de \mathcal{E} denotada por $H(\mathcal{E})$ é:

$$H(\mathcal{E}) = - \sum_{i \in I} P(E_i) \log P(E_i).$$

A entropia conjunta de dois sistemas \mathcal{E} e \mathcal{F} é

$$H(\mathcal{E} \wedge \mathcal{F}) = - \sum_{i \in I} \sum_{j \in J} P(E_i \cap F_j) \log P(E_i \cap F_j)$$

Entropia

A entropia condicional de \mathcal{E} sobre \mathcal{F} é:

$$H(\mathcal{E}|\mathcal{F}) = \sum_{i \in I} \sum_{j \in J} P(E_i \cap F_j) I(E_i, F_j) \quad (1)$$

$$= \sum_{i \in I} \sum_{j \in J} P(E_i \cap F_j) \log \frac{P(E_i \cap F_j)}{P(F_j)} \quad (2)$$

Observações

- $H(\mathcal{E})$ é a informação média das informações de \mathcal{E} .
- Medida da desordem.

Exemplo

Seja:

$$X = \begin{cases} 1 & \text{prob. } p \\ 0 & \text{prob. } 1 - p \end{cases}$$

Determine $H(X)$.

Exemplo

Seja:

$$X = \begin{cases} a & \text{prob. } 1/2 \\ b & \text{prob. } 1/4 \\ c & \text{prob. } 1/8 \\ d & \text{prob. } 1/8 \end{cases}$$

Determine $H(X)$.

Nota: Neste e nos demais exemplos, considere X como sendo o sistema de eventos formado pelos eventos do conjunto X e suas probabilidades associadas.

Exemplo

y/x	1	2	3	4
1	1/8	1/16	1/32	1/32
2	1/16	1/8	1/32	1/32
3	1/16	1/16	1/16	1/16
4	1/4	0	0	0

Determine $H(X \wedge Y)$, $H(X)$, $H(Y)$ e $H(X|Y)$.

Exemplo

Uma moeda honesta é lançada até que a primeira cara aparece. Seja X o número de tentativas, determine $H(X)$.

Teorema da Entropia

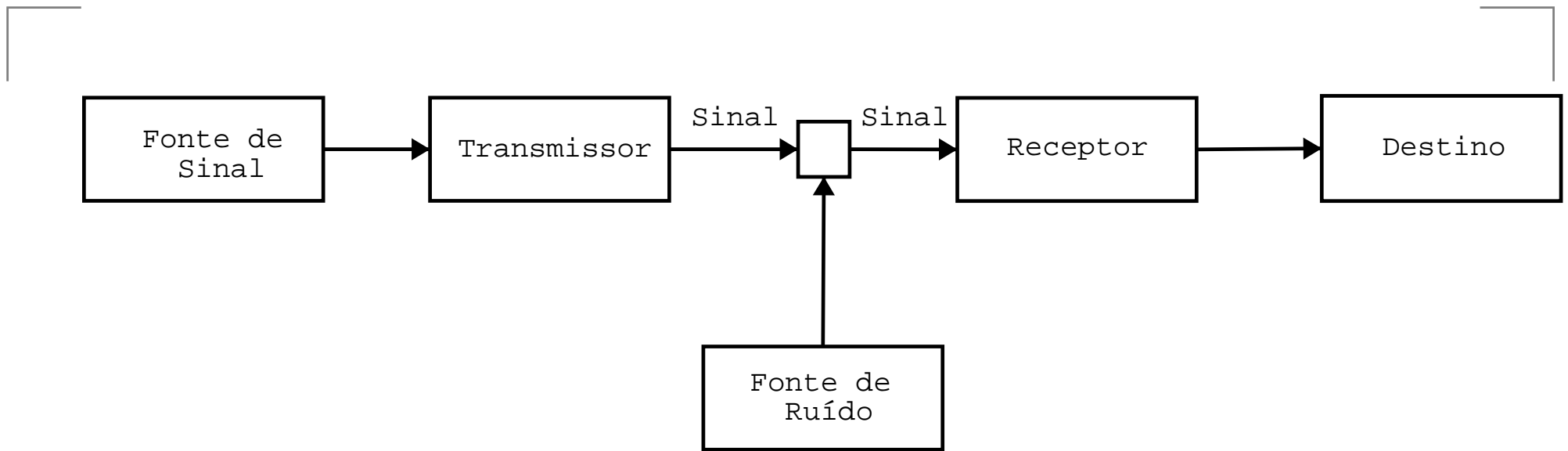
Suponha que $\mathcal{E} = [E_i, i \in I]$ é um EPF. Então $0 \leq H(\mathcal{E}) \leq \log |I|$. A igualdade no extremo inferior ocorre se e somente se todos eventos menos um ocorre com probabilidade 1. A igualdade com o extremo superior ocorre se todos os eventos de \mathcal{E} são equiprováveis.

Informação e Entropia

- $I(\mathcal{E}, \mathcal{F}) = H(\mathcal{E}) + H(\mathcal{F}) - H(\mathcal{E} \wedge \mathcal{F})$
- $H(\mathcal{E}|\mathcal{F}) = H(\mathcal{E} \wedge \mathcal{F}) - H(\mathcal{F})$

Canal e Capacidade do Canal

Definição do Canal



Definição do Canal

Fonte Gera Sinal.

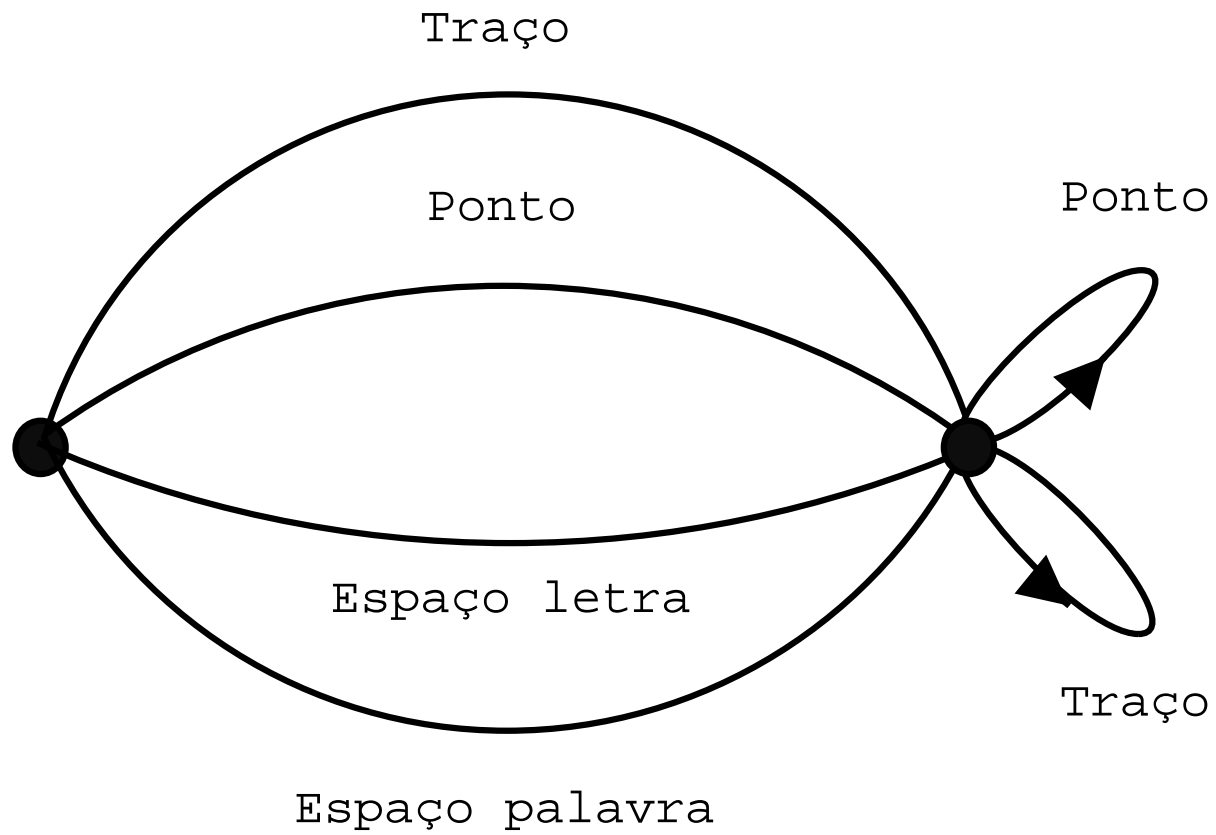
Transmissor Transforma o sinal em um sinal adequado para ser transmitido.

Canal Meio Físico.

Receptor Mecanismo responsável por receber o sinal.

Destino Pessoa ou máquina que recebe o sinal.

Telégrafo



Classificação Canais

- Canais Discretos, Contínuos ou Mistos
- Canais com ruído ou sem ruído.

Especificação do Problema

- O canal é discreto.
- O canal receberá sinais formados por letras de um alfabeto com n letras, $\{a_1, \dots, a_n\}$.
- O canal responderá com sinais formados com letras de um alfabeto com m letras, $\{b_1, \dots, b_r\}$.
- Os canais possuem ruído.

Definições

O conjunto de todas as palavras de tamanho k do alfabeto A .

$$A^k = \underbrace{A \times A \times \dots \times A}_{k \text{ vezes}}$$

O conjunto de todas as palavras não vazias de A

$$A^+ = \bigcup_{k=1}^{\infty} A^k$$

Canais sem Memória

A probabilidade de b_j ser a saída se a_i for a entrada não depende do tempo, nem da história para todos os i e j .

Probabilidades de transição

Sejam $A = \{a_1, \dots, a_n\}$, $B = \{b_1, \dots, b_r\}$, $a_i \xrightarrow{q_{ij}} b_j$,
 $Q = [q_{ij}]$ uma matriz $n \times r$ e

$$\sum_{j=1}^r q_{ij} = 1.$$

Canal Binário e Simétrico

Um canal é binário simétrico é um canal sem memória com $A = \{0, 1\}$, $B = \{0, 1\}$. A probabilidade de que um bit seja transmitido corretamente é p .

$$Q = \begin{bmatrix} q_{00} & q_{01} \\ q_{10} & q_{11} \end{bmatrix} = \begin{bmatrix} p & 1 - p \\ 1 - p & p \end{bmatrix}$$

Observação

A transmissão de n bits é equivalente a n tentativas de Bernoulli com probabilidade de sucesso p . A probabilidade de k sucessos é $C_{n,k}p^k(1-p)^{(n-k)}$ e o número esperado de sucessos é np .

Exercício

Para um canal sem memória nós temos $A = \{0, 1\}$, $B = \{0, 1, *\}$ e o canal trata de forma simétrica os dígitos de entrada, cada dígito possui probabilidade p de ser transmitido corretamente, probabilidade q de ser trocado e probabilidade r de ser transformado em $*$. Note que $p + q + r = 1$.

1. Escreva a matriz de transição de probabilidades em termos de p , q e r .
2. Em termos de n , p e k , qual é a probabilidade de ocorrerem exatamente k erros na transmissão de uma palavra de tamanho n neste canal?

Exercício - Continuação

1. Suponha que $*$ seja eliminado do alfabeto de saída usando um lançamento de moeda com uma moeda honesta. Sempre que $*$ é recebido a moeda é lançada se obtemos cara então $*$ é lido como 0, caso contrario ele é lido como 1. Como é a nova matriz de probabilidades? O novo canal é binário e simétrico?
2. Suponha que $*$ seja eliminado do alfabeto de saída misturando-o com 1. Sempre que $*$ é recebido lemos 1. Qual é a matriz de probabilidades? O novo canal é simétrico?

Exercício

Um canal binário e simétrico possui confiabilidade p .

1. Qual é o valor mínimo de p para que tenhamos pelo menos 95% de chance de não ocorrer erros na transmissão de uma palavra de comprimento 15?
2. Obtenha a desigualdade que p deve satisfazer se exigimos que até um erro será tolerado.
3. Qual é o mínimo valor de p permitido se o número médio de erros na transmissão de uma palavra binária de comprimento 15 não deve ser maior que 1/2?

Transmissão de Mensagens

Uma letra do alfabeto de entrada A é enviada pelo canal e uma letra do alfabeto B é recebida.

Esta situação pode ser descrita em termos de um Espaço de Probabilidades Finito.

Frequências de Entrada

- $A = a_1, \dots, a_n$

- $p_i =$ frequência relativa de transmissão

$$p_i = \frac{\text{número de vezes que aparece } a_i}{\text{número de caracteres enviados}}$$

- $S = \{(a_i, b_j), i \in \{1, 2, 3, \dots, n\}, j \in \{1, 2, \dots, r\}\}$

Resultados

- $P(a_i) = p_i$
- $P(b_j|a_i) = q_{ij}$
- $P(a_i, b_j) = p_i q_{ij}$
- $P(b_j) = \sum_{k=1}^r p_k q_{kj}$

Exercício

Suponha que uma fonte emite três mensagens: M_1 , M_2 e M_3 , com frequências de 30%, 50% e 20% respectivamente. Estas mensagens devem ser transmitidas por um canal binário. Com este objetivo as mensagens são codificadas como $M_1 = 11111$, $M_2 = 1000001$ e $M_3 = 1100$. Encontre as frequências de entrada de 0 e 1.

Exercício

Um canal binário e simétrico com confiabilidade p é usado em uma tarefa de comunicação para a qual as frequências de entrada dos caracteres 0 e 1 são respectivamente $2/3$ e $1/3$. Encontre as frequências de saída dos caracteres 0 e 1 em função de p .

Exercício

Seja $A = \{a_1, a_2, a_3\}$ e $B = \{b_1, b_2, b_3\}$,

$$Q = \begin{bmatrix} .94 & .04 & .02 \\ .01 & .93 & .06 \\ .03 & .04 & .93 \end{bmatrix},$$

$p_1 = .4$, $p_2 = .5$ e $p_3 = .1$. Encontre as frequências de saída $P(b_1)$, $P(b_2)$ e $P(b_3)$.

Exercícios

Considere um canal sem memória nós temos $A = 0, 1$ e $B = 0, 1, *$ e o canal trata de forma simétrica os dígitos de entrada, cada dígito tem probabilidade p de ser transmitido corretamente, q de ser transformado em outro dígito e r de ser confundido, de forma que a saída seja $*$. Observe que $p + q + r = 1$.

- Encontre a matriz de transição de probabilidades em termos de p , q e r .
- Determine a probabilidade de ocorrer k erros.

Exercícios - Continuação

- Suponha que * seja eliminado do alfabeto de saída. Sempre que * seja recebido 0 ou 1 é escolhido. Determine a nova matriz de probabilidades. O canal agora é binário e simétrico?
- Suponha agora que * seja eliminado do alfabeto de saída. Determine a nova matriz de transição. O canal é binário e simétrico?

Minimização de Funções de Várias Variáveis

Seja $F = F(x_1, \dots, x_n)$ uma função definida em um domínio A , por simplicidade assumimos que F seja diferenciável. Os pontos de extremo desta função podem ser obtidos através das equações:

$$\frac{\partial F}{\partial x_k} = 0$$

Funções sujeitas a vínculos

Método dos Multiplicadores de Lagrange Considere a função $F = F(x_1, \dots, x_n)$ sujeito ao vínculo $G(x_1, \dots, x_n) = 0$. Para determinar os extremos desta função inicialmente resolvemos o sistema:

$$\frac{\partial F}{\partial x_k} - \lambda \frac{\partial G}{\partial x_k} = 0$$

$$G(x_1, \dots, x_n) = 0$$

Exemplo

Um fazendeiro possui P metros de cerca e quer cercar a maior área retangular possível. Como deve escolher os lados do retângulo?

Descrição do Modelo

- Os símbolos a serem transmitidos são variáveis aleatórias independentes.
- O canal é ruidoso mas não possui memória.

Capacidade do Canal

Qual é a capacidade máxima de transmissão de um dado canal?

Como o canal é definido pela natureza e não podemos alterá-lo, para atingir esta taxa máxima de transmissão devemos mexer nas p_i , ou seja nas frequências de entrada do canal.

Que valores p_i deve assumir para que possamos maximizar a taxa de transmissão de dados?

Maximizar a informação mútua

Considere o EPF $A = \{\text{enviar a letra } a_i\}$ e $B = \{\text{receber a letra } b_j\}$. A capacidade máxima do canal ocorre quando $I(A, B)$ é máxima.

$$\begin{aligned} I(A, B) &= \sum_{i=1}^n \sum_{j=1}^r P(a_i \cap b_j) \log \frac{P(a_i \cap b_j)}{P(a_i)P(b_j)} \\ &= \sum_{i=1}^n \sum_{j=1}^r p_i q_{ij} \log \frac{p_i q_{ij}}{\sum_{t=1}^n p_t q_{tj}} \end{aligned}$$

Nota: Esta hipótese será justificada *a posteriori*.

Determinação das p_i

Matematicamente este problema consiste em determinar as freq. p_i sujeitas ao vínculo:

$$\sum_{i=1}^n p_i = 1.$$

Para realizarmos este feito vamos utilizar a técnica dos multiplicadores de Lagrange.

Resumo da ópera

O valor máximo de $I(A, B)$ assume é dado por:

$$I'(A, B) = C = \sum_{j=1}^r q_{kj} \log \frac{q_{kj}}{\sum_{t=1}^r p_t q_{tj}}$$

C é a capacidade do canal e está relacionada a taxa máxima possível de transmitir dados através do canal.

Teorema

Suponha um canal sem memória com alfabeto de entrada $A = \{a_1, \dots, a_n\}$ e alfabeto de saída $B = \{b_1, \dots, b_r\}$ e probabilidades de transição q_{ij} . Existem freqüências ótimas para este canal. Se p_1, \dots, p_n são números positivos e reais, estas freqüências de entrada ótimas se e somente se temos:

$$C = \sum_{j=1}^r q_{kj} \log \frac{q_{kj}}{\sum_{t=1}^r p_t q_{tj}}$$

e além disso $I'(A, B) = C$.

Interpretação de Shannon

Seja r a taxa de entrada de caracteres, $rH(A)$ é a taxa de transmissão de informação. Qual é a taxa de informação recebida pelo receptor:

$$rI(A, B) = r(H(A) - H(A|B))$$

$$H(A|B) = \sum_{ij} P(a_i \cap b_j) \log \frac{P(a_i \cap b_j)}{P(b_j)} \quad (3)$$

- Se $P(a_i \cap b_j) = P(b_j)$ (os caracteres são enviados sem erro) $H(A|B) = 0$.
- Se a_i e b_j são independentes (nenhuma transmissão pode ser transmitida e $H(A|B) = H(A)$ e $I(A, B) = 0$.

Observação

O termo $H(A|B)$ corrige o fato que quando temos a_i e b_j des-correlacionados ainda assim teríamos casos em que os caracteres seriam corretamente decodificados por pura sorte!
A $H(A|B)$ é chamada também de equivocação.

Canal binário e simétrico

$$C = p \log_2 p + (1 - p) \log_2 (1 - p)$$

Exemplo:

$$p=0.99 \quad C=0.919$$

Exercícios

Verifique diretamente que $f(p) = p \log_2 p + (1 - p) \log_2 (1 - p)$ atinge seu valor máximo em 0 e em 1 e o seu mínimo em $1/2$.

Exercício

Considere $A = B = \{0, 1\}$ e um canal não simétrico, suponha que 0 possua probabilidade p de ser recebido 0 e 1 possua probabilidade q de ser recebido 1. Sejam p_0 e p_1 as freqüências de entrada. Em termos de p , q , p_0 e p_1 escreva $I(A, B)$ e determine a capacidade do canal.

Exercícios

Considere um canal sem memória nós temos $A = \{0, 1\}$ e $B = \{0, 1, *\}$ e o canal trata de forma simétrica os dígitos de entrada, cada dígito tem probabilidade p de ser transmitido corretamente, q de ser transformado em outro dígito e r de ser confundido, de forma que a saída seja $*$. Observe que $p + q + r = 1$. Encontre a capacidade do canal e as freqüências de entrada ótimas.

Exercício

Suponha que $A = B = \{a_i, \dots, a_n\}$ e que o canal é perfeitamente confiável. Encontre a capacidade do canal e as frequências ótimas.

Equipartição Assintótica

Seja X uma variável aleatória que pode assumir k valores e $p(x_1, x_2, \dots, x_n)$ a probabilidade que uma variável aleatória X tenha assumido ao longo do tempo os valores x_1, x_2, \dots, x_n . Para eventos típicos temos que:

$$\frac{1}{n} \log_2 \frac{1}{p(x_1, x_2, \dots, x_n)} = H.$$

Isto implica que $N_{TIP} = 2^{nH}$.

Idéia de Compressão

Usar $nH + 2$ bits para designar os elementos de \mathcal{H} (conjunto das seqüências típicas) e usar $n \log K + 2$ bits para designar os demais. Neste caso temos que $\langle l \rangle \sim H$.