**第 3 次课 学时 2**

|  |  |
| --- | --- |
| **授课章节内容** | 第三章单符号离散信道 |
|  | 3.1信道的数学模型 |
| 3.2信道的交互信息量与条件互信息量 |
| 3.3 平均交互信息量及其性质 |
| 3.4 信道容量及其一般算法 |
| **教学目标** | 教学目标2 |
| **支撑毕业要求** | 毕业要求1-3 |

**教学要求：**

**教学要求：**

1. 知识目标

* 了解信道的数学模型及其分类；掌握平均互信息概念及性质，掌握信道容量的计算

2. 能力目标

* 具备计算信道平均互信息的能力。

1. 素质目标

* 激发学生能结合所学知识探索信息领域的热情，结合思政课程建设，介绍我国通信领域专家在该领域的贡献，培养家国情怀。

**教学重点与难点**：

* 平均互信息与信道容量的计算

**教学过程设计：**

* 由单符号事件间的互信息量引出平均互信息的概念

**教学方法及手段：**

PPT为主，例题 板书为辅

| **讲授与指导内容** | | **讲课、互动内容设计** |
| --- | --- | --- |
| 第3章 单符号离散信道  信息论中有关信道问题的讨论，使用编码信道模型，即用信道输入输出符号转移概率或转移概率密度函数表征信道特性。  如果信道输入和输出的是随机过程，则对应的是波形信道；如果信道的输入和输出是随机矢量，而且矢量中每个随机变量的取值是连续的，则对应的信道是连续信道，每个随机变量的取值是离散的，则对应的信道是离散信道。离散信道中，若输入输出分别仅有一个随机变量，这种信道称为单符号离散信道。  本章将对单符号离散信道的信息传输、信道容量计算等一系列基本问题展开谈论。  3.1信道的数学模型  单符号离散信道模型如图3.1所示。    图3.1单符号离散信道模型  输入变量*X*有*r*种取值，即输入符号集*X:{a1,a2,...,ar}*，输出变量*Y*有*s*种取值，即输出符号集{*b1,b2,...bs*}。信道转移概率    共有*r×s*个取值，体现了信道的符号传递特性。  写成矩阵形式，形成一个（*r×s*）阶矩阵，矩阵行数代表信道输入符号个数，矩阵列数代表输出符号个数。  而且满足  ，表明信道矩阵行和是1。信道特性也可以形象、直观地用信道转移图表示    图3.2 信道转移图  【例3.1】二进制对称信道，简记为BSC（Binary Semmetric Channel），输入输出符号集分别为*X*:{0,1}，*Y*:{0,1}，信道转移概率满足P*Y|X*(0/0)=P*Y|X*(1/1)=1-*p*，P*Y|X*(1/0)=P*Y|X*(0/1)=*p*，*p*为错误传输概率。写出信道的转移概率矩阵，并画出转移概率图。  解：转移概率矩阵为，相应的信道转移概率矩阵如图3.3所示    图3.3 二进制对称信道转移概率图  【例3.2】二进制删除信道，简记为BEC（Binary Erasure Channel）其中输入集*X*:{0,1}，输出集*Y*:{0,2,1}转移概率如图    图3.4 二进制删除信道转移概率图  写出信道的转移概率矩阵。  解：信道的转移概率矩阵为  3.2 信道的交互信息量  在求解信道输入输出单个符号对应互信息量前，先讨论信道输入输出变量的统计特性。通信系统模型如图3.5所示    图3.5 系统模型  为方便讨论，首先定义几种概率的名称。  1.先验概率  信源*X*输出符号*ai*的概率*p*(*x*=*ai*)=*p*(*ai*)称为先验概率。  2.正向转移概率  从信道输入符号*ai*到信道输出符号*bj*的条件概率  （3.1）  称为正向转移概率。  3.后验概率  从信道输出符号bj到输出符号ai的条件概率  （3.2）  称为后验概率，又称之为反向转移概率。  利用，则有  （3.3）  看到，已知先验概率和信道转移概率、后验概率即为确定。  4.联合概率  信源符号ai和信道输出符号bj的联合概率  （3.4）  由信源概率和信道转移概率唯一确定。  类似信道转移概率矩阵可以写出联合概率矩阵，也是r行s列矩阵。  （3.5）  5.信宿概率  信道输出符号概率，已知信源符号概率和信道转移概率，即可求得信道输出符号概率。  现在转入互信息量的讨论。  在通信系统中，信源发出某符号ai,由于受到噪声的随机干扰，在信道的输出端输出符号*ai*的某种变型*bj*，按信息的定义，信宿收到*bj*后，从*bj*中获取关于*ai*的信息量I(*ai*; *bj*)，等于信宿收到bj前、后，对符号ai的不确定性的消除，即有  （3.6）  信宿收到*bj*前，对信源发符号*ai*的先验不确定性  （3.7）  信宿收到*bj*后，对信源发出的符号*ai*的后验不确定性  （3.8）  则可得互信息为  （3.9）  信宿收到*bj*后，从*bj*中获取关于*ai*的信息量I(*ai*; *bj*)称为输入符号*ai*和输出符号*bj*之间的交互信息量，简称为互信息。它表示信道在把输入符号*ai*传递为输出符号*bj*的过程中，信道所传递的信息量。（3.9）式称为符号*ai*和*bj*之间的互信息函数。  现在就互信息量表达式所代表的物理含义做进一步说明。  1.当时，有  （3.10）  表明收到的*bj*(*j*=1,2,...,*s*)后，即可确切无误地判断发端符号为*ai*，消除对*ai*的全部不确定性，收端获得关于*ai*的全部信息量I(*ai*) (*i*=1,2,...,*r*)。  2.当时，有  （3.11）  这就意味着，收信者收到bj后，判断信源发*ai* (*i*=1,2,...,*r*)的可能性，比对于收到*bj* (*j*=1,2,...,*s*)前判断信源发*ai* (*i*=1,2,...,*r*)的可能性更大；也就是说收到*bj* (*j*=1,2,...,*s*)后对信源发*ai* (*i*=1,2,...,*r*)的不确定性，比收到*bj* (*j*=1,2,...,*s*)前有所减小，收信者从*bj* (*j*=1,2,...,*s*)中就可获取关于*ai* (*i*=1,2,...,*r*)的一定信息量。  3.当时，有  （3.12）  由得到，符号*ai*与符号*bj*统计独立。  说明，收信者收到*bj* (*j*=1,2,...,*s*)前、后，判断信源发*ai* (*i*=1,2,...,*r*)的可能性大小没有任何变化，收信者在收到*bj* (*j*=1,2,...,*s*)前、后，对判断信源发*ai* (*i*=1,2,...,*r*)的不确定性没有任何减小，收信者从*bj* (*j*=1,2,...,*s*)中得不到关于*ai* (*i*=1,2,...,*r*)的任何信息量。  4.当时，有  （3.13）  表明，由于信道噪声的干扰，收到符号*bj* (*j*=1,2,...,*s*)后，猜测符号*ai* (*i*=1,2,...,*r*)的难度反而加大。  互信息量*I*(*ai*; *bj*)的另外两种表达形式  1.第一种表达形式  （3.14）  其含义是：通信前，条件*ai*与*bj*统计独立。即，表达联合条件*I*(*ai*; *bj*)不确定性；通信后，条件*ai*与条件*bj*建立统计关联，表达通信后，联合条件*I*(*ai*; *bj*)的确定性。二者差，同样表明，信宿收到*bj*后，从*bj*中获取关于*ai*的信息量*I*(*ai*; *bj*)等于通信前后不确定性的消除。  2.第二种表达形式  （3.15）  ，构成正向信道；  ，构成反向信道。  说明在反向信道中，从*ai*中获得关于*bj*的信息量等于在正向信道中从*bj*中获得关于*ai*的信息量，即*I*(*ai*; *bj*)= *I*(*bj*; *ai*)。  【例3.3】二元删除信道，简记为BEC（Binary Erasure Channel）：其中输入符号集*A*={*a*0,*a*1}={0,1}，输出符号集*B*={*b*0,*b*1,*b*2}={0,1,2}，*p*(*a*0)=p(*a*1)=0.5,写出信道的转移概率矩阵及互信息量*I*(*ai*; *bj*)。  解：设信道的转移概率矩阵为    联合概率为    则*P*(b*j*)为    那么互信息量*I*(*ai*; *bj*)用矩阵形式简写为    【例3.4】4个等概率消息，编程的码字为M1=000 ,M2=011,M3=101,M4=110，通过如图所示二元对称无记忆信道(ε<0.5)传输，求：①事件“接收到第一个数字为0”与发送M1之间的互信息；②当接收到第二个数字也为0时，关于M1的附加信息；③当接收到第三个数字也为0时，又增加了多少关于M1的信息？    图3.6 二元对称无记忆信道  解：记“0”表示第一个接收数字为0，“00”表示第一、二个接收数字都为0，“000”表示前三个接收数字都为0；表示接收符号的概率；*p(y*/*x*)为信道的转移概率。  ➀；  互信息：。  ➁互信息：  附加信息：  ➂  互信息：  又增加的信息为  3.3 条件互信息量  现在把目光转到级联信道的交互信息的讨论上。  如图3.7所示信道I与信道II串接。信道I的输入符号集X={a1,a2,...,a*r*}，输出符号集Y={b1,b2,...,b*s*}，信道II的输入符号集Y={b1,b2,...,b*s*}，输出符号集Z={c1,c2,...,c*t*}    图3.7 信道I与信道II串接  信道I的传递概率：    信道II的传递概率：    则  由X、Y、Z的联合概率*p*(*aibjck*)可以求得其他各种概率分布。  一维分布：        二维分布：        进一步求得条件概率分布：              在串接信道输出某符号c*k*条件下，从符号b*j*中获取符号a*i*的信息量定义为  （3.16）  该式表明，c*k*已知条件下，b*j*出现前后对a*i*的条件不确定性的消除。  用p(b*j*c*k*)同时乘上式右边的分子、分母，则有：  （3.17）  上式表明，I信道通信前后，输入输出端同时出现a*i*和b*j*的条件不确定性的消除。  由3.17式，做进一步变换，有    所以有  在图3.7中，随机变量c*k*与随机变量*X*和*Y*的联合符号(*aibj*)之间的相关交互信息量    这表明，相关交互信息量等于b*j*与c*k*之间的交互信息量，再加上b*j*已知的条件下，a*i*与c*k*之间的条件交互信息量所得之和。同样    这表明，相关交互信息量也等于a*i*与c*k*之间的交互信息量，再加上a*i*已知的条件下，b*j*与c*k*之间的条件交互信息量所得之和。  【**例3.5**】如下表所示，列出了无失真信源编码的信源消息、消息的先验概率以及每一个消息所对应的码字。  表3.1   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 信源消息 |  |  |  |  |  |  |  |  | | 码字 | 000 | 001 | 010 | 011 | 100 | 101 | 110 | 111 | | 消息概率 | 1/4 | 1/4 | 1/8 | 1/8 | 1/16 | 1/16 | 1/16 | 1/16 |   试以消息a*5*及相应码字（100）为例，分别说明码字（100）中每一个码符号对消息a*5*提供的信息量。  解：根据相关交互信息量的理论可有    以下分别计算其中各项条件互信息。  （1）  其中，p(a*5*/1)表示收到码符号“1”后，判断信源发消息a*5*的后验概率。因为收到码符号“1”后，再收到码符号序列“00”就构成码字（100），即消息a*5*出现。所以后验概率p(a*5*/1)= p(00/1)，即有p(a*5*/1) = p(00/1)= p(100)/ p(1)。  其中，码字（100）出现的概率p(100)等于消息a*5*出现的概率，即有p(100)=p(a*5*)=1/16  从表3.1中可看出，8个码字中有4个码字（100）、（101）、（110）、（111）的第一个码符号是“1”，所以有    即可得  这样就有  （2）  其中，表示收到码符号“10”后，判断信源发消息的后验概率。因为收到码符号“10”后，再收到码符号序列“0”就构成码字（100），即消息出现。所以后验概率，即有  从表3.1中可看出，8个码字中有2个码字（100）、（101）的前两个码符号序列是“10”，所以有    即可得  这样就有  （3）  其中，表示收到码符字（100）后，判断信源发消息的后验概率。显然收到（100）后也就是收到了消息，所以有  这样就有  以上三项计算结果表明，在消息相对应的码字（100）中：第一个码符号“1”提供关于消息的信息量；在收到第一个码符号“1”的条件下，第二个码符号“0”提供关于的条件交互信息量；在收到第一个码符号“1”和第二个码符号“0”组成的码符号序列“10”的条件下，第三个码符号“0”提供关于的条件交互信息量。所以，从码字（100）中的三个码符号总共提供关于消息的相关交互信息量    另一方面，消息的自信息量  (比特)  这从信息测量的角度证实了由于消息与相应的码字（100）是一一对应的确定关系，相关交互信息量就是消息的自信息量。  3.4 平均交互信息量  *I(ai,bj)*表示单个事件间的交互信息量，要求信道两端平均一对符号传递信息量的多少，就要计算平均交互信息量，如图3.8和图3.9所示    图3.8 信息传输方向为X到Y 图3.9 信息传输方向为Y到X  首先给出两种特殊形式  1.  （3.18）  这表示信宿收到b*j*后，获得有关信源的信息量。  2.  （3.19）  这表示信宿收到a*i*后，获得有关信源的信息量。  由上述两个式子可知，*X*与*Y*间平均传递一个符号所传输的平均信息量*I(X,Y)*应该是*I*(a*i*,b*j*)在联合集*XY*中的统计均值，即  （3.20）  同样，*I(X;Y)*也有三种不同的表达形式  1.用信源概率*p(ai)*和正向传输概率*p(bj/ai)*表达  （3.21）  其中，  （3.22）  表示在随机变量的前提下，对随机变量*X*仍然存在的平均不确定性。而条件熵  （3.23）  表示收到随机变量*Y*后，对随机变量*X*仍然存在的平均不确定性，通常称它为疑义度。  （3.21）式表明，从收到*Y*中获取关于*X*的平均交互信息量，等于收到*Y*前对*X*的平均不确定性，与收到*Y*后对*X*仍然存在的平均不确定性之差，即收到*Y*前、后，关于*X*的平均不确定性的消除。  2.用信宿概率和反向传输概率表达  （3.24）  其中，  （3.25）  表示在随机变量的前提下，随机变量Y仍然存在的平均不确定性。而条件熵  （3.26）  表示表示在反向信道中，收到随机变量X后，对随机变量Y仍然存在的平均不确定性。这个“反向疑义度”一般称为噪声熵。  （3.24）式表明，对于反向信道来说，从输出随机变量X中，获取关于Y的平均交互信息量，等于信宿收到X前，对Y的先验不确定性，与信宿收到X后，对Y仍然存在的后验平均不确定性之差，即通信前、后，关于Y的平均不确定性的消除。  3.用信源概率、信宿概率和联合概率表达  （3.27）  其中，    表示通信后，信道两端同时出现X和Y的后验平均不确定性，通常称它为共熵。  （3.27）式表明，信源Ｘ通过传递概率为的信道输出随机变量Y，信道传递的平均交互信息量，等于通信前（随机变量X和Y统计独立）随机变量X和Y同时出现的平均不确定性，与通信后（随机变量X和Y由信道传递概率相联系）信道两端同时出现随机变量X和Y的平均不确定性之差，即通信前、后，随机变量X和Y同时出现的平均不确定性的消除。  上述讨论的通信系统中各类熵的关系可用维拉图形象表示，如图3.10所示。    图3.10通信系统中各类商关系图  【**例3.6**】设信源*X*的符号集，先验概率分布为；。信道的输入符号集，输出符号集：，传递概率。现将信源X与如图3.11信道相接。  （一）试写出平均交互信息量的一般表达式；  （二）若，；。试计算如图3.12所示反向信道的。    图3.11 正向信道传递概率图 图3.12反向信道传递概率图  解：（一）平均交互信息量的一般表达式  （1）各联合概率          （2）随机变量*Y*的概率分布      （3）求随机变量*Y*的熵    （4）由求得条件熵    （5）求得平均交互信息量    这说明平均交互信息量是信源概率分布和信道传递概率的函数。  （二）计算反向信道的  （1）各联合概率          （2）随机变量*Y*的概率分布      （3）求随机变量*Y*的熵    （4）由，求得噪声熵    （5）求得平均交互信息量    3.5 平均交互信息量的性质  如上节所述，平均交互信息量*I(X;Y)*除具有对称性以外，即*I(X;Y)= I(Y;X)*，还具有如下基本性质。  1.平均互信息的非负性    当且仅当*X*和*Y*统计独立时，等式成立。  **【证明】**利用詹森不等式得    即有    当且仅当对一切i，j都有  即当*X和Y*统计独立时，  2.平均互信息的极值性  由上述性质，直接得到      **【证明】**由于，而是对求统计平均，即    因此有    同理    所以      即    3.平均互信息的凸函数性  **定理3.1** 信道两端随机变量*X*和*Y*之间的平均互信息量*I(X;Y)*，在信道转移概率p(bj/ai)给定条件下，是输入随机变量*X*的概率分布的∩型凸函数。  **【证明】**当条件概率P(y/x)是固定时，平均互信息*I(X;Y)*只是P(*x*)的函数。简写成。现选择输入信源*X*的两种已知的概率分布和。其对应的联合概率分布为和，因而平均互信息分别为和。再选择输入变量X的另一种概率分布，令，和，而，因而得其相应的平均互信息为。  根据平均互信息的定义得    上式中根据概率关系    所以得    因为log*x*是*x*的∩型函数，所以对上式中第一项，根据詹森不等式得    同理    又因为和都小于1且大于0，即得，因而    因此根据凸函数的定义知，*I(X;Y)*是概率分布的∩型凸函数。  **定理3.2** 信道两端随机变量X和Y之间的平均交互信息量*I(X;Y)*，在信源概率分布给定的条件下，是信道转移概率的∪型凸函数。  **【证明】**当概率分布固定时，平均互信息*I(X;Y)*只是条件概率的函数，简写成。选择两种条件概率分别为和。相对应的平均互信息分别为和，再选择第三种条件概率满足。设相应的平均互信息为。其中,和。因而求得    运用詹森不等式，上式中第一项为    同理    所以，即    根据凸函数的定义得：平均互信息是条件概率的∪型凸函数。  3.6 信道容量及其一般算法  3.6.1 信道容量的定义  信道的信息传输率定义为信道中平均每个符号所传送的信息量，即平均互信息。    信道的信息传输速率定义为信道平均每秒传输的信息量。若传输一个符号平均需要t秒，则信道的信息传输速率表示为    给定某个信道，平均交互信息量是信源概率分布的上凸型函数，存在极大值，这个极大值就定义为信道容量。    信道的最大信息传输速率是信道容量的另一种表述形式    3.6.2 信道容量的一般算法  给出一个典型例题  【**例3.7**】设离散信道的输入符号集为，输出符号集为，其信道转移概率矩阵为,求其信道容量及最佳输入分布。  解：列得    解得    则信道的信道容量  由得  由于可写成下式    因此    因此，最佳分布为    现在验证以上结果    同理可以计算        显然，    而每个符号贡献的互信息也正好是前文求解出的信道容量，证实了该求解过程是正确的。 | | 问:信道矩阵的行数和列数各具有怎样的含义？  二进制对称信道是最简单的信道，同时也是理论和实践中最重要的信道  强调二进制删除信道的工程背景。  渗透这样的一个概念：互信息本质是基于信息的定义  补充：板书给出信源，信宿之间的矩阵形式依赖关系  平均互信的三种表达方式是本课程的核心概念之一，要求熟练掌握  解释互信息极值性的物理含义  平均互信息的凸性是计算信道容量和信源有损压缩的理论基础，非常重要  信道容量的一般算法不需要掌握，了解即可，请参阅教材 |
| **复习要点** | 能够正确解释平均互信息概念；能够计算信道容量 | |
| **思考题** | 1.平均互信息的量纲是什么？  2. 信道容量计算有什么实际意义？ | |
| **作业题** | 3.1 3.2 | |

作者签名：

****