信息论基础教程习题解答与实验指导

李　梅　李亦农　编著

北京邮电大学出版社
·北京·

内 容 简 介

本书为《信息论基础教程》的配套教材，包括原书各章的习题解答、新增的习题及解答以及上机实验的参考程序等。

本书作为教材配套的实验指导，强调实践环节，编写了5个实验的实验指导：信道容量的迭代算法、唯一可译码判决准则、Huffman编码、LZW编码和Shannon编码，每个实验都给出了实验目的、实验要求、算法和参考程序。

图书在版编目(CIP)数据

信息论基础教程习题解答与实验指导/李梅，李亦农编著 .—北京：北京邮电大学出版社，2005

ISBN 7-5635-1163-6

Ⅰ.信... Ⅱ.①李...②李... Ⅲ.信息论—高等学校—教学参考资料 Ⅳ.G201

中国版本图书馆CIP数据核字(2005)第120215号

书　　名：信息论基础教程习题解答与实验指导

编　　著：李梅　李亦农

责任编辑：李欣一

出 版 者：北京邮电大学出版社(北京市海淀区西土城路10号)

邮编：100876　电话：62282185　62283578(Fax)

电子信箱：publish @ bupt.edu.cn

经　　销：各地新华书店

印　　刷：北京源海印刷有限责任公司

印　　数：1—5 000册

开　　本：787 mm×1 092 mm　1/16　　印张：9.5　　字数：206千字

版　　次：2005年10月第1版　2005年10月第1次印刷

ISBN 7-5635-1163-6/TN·401　　定　价：15.00元

·如有印装质量问题，请与北京邮电大学出版社发行部联系·

前言

随着信息技术的飞速发展，信息论已经成为越来越多的学校和专业的专业基础课和专业选修课，而同时各个学校都在大量压缩专业课课时，一般通信工程和电子信息专业信息论的授课时数都在34课时左右，因此迫切需要一本内容精简、符合少学时教学要求的教材。基于这样的考虑，我们在借鉴了国内外众多的信息论优秀教材和参考书之后编写了《信息论基础教程》教材。教材以香农的3个编码定理为中心，重点讲述了相关的基本概念、基本原理和基本方法。

本书为《信息论基础教程》的配套教材，包括原书各章的习题解答、新增的习题及解答以及上机实验的参考程序等。

习题是读者深入理解和消化基本理论，锻炼独立解决问题能力的重要环节。本书给出了教材的习题解答，并且补充了一些典型习题，所选配的习题是从国内外大量优秀习题中精选而来的，具有一定的代表性，同时很多题与日常生活密切相关，比较新颖，具有一定的趣味性。大部分习题解答需要综合运用所学知识，能有效地帮助学生巩固、加深对授课内容的理解。本书的解答仅仅作为读者的学习参考，希望读者能通过自己的思考得出不同的解题方法。

信息论是一门理论和实践紧密结合的课程，如果仅仅停留于理论，学生学到的就是死的知识，不能很好地调动学生对课程学习的兴趣，很难体现培养创新思维和动手能力的要求。本书作为教材配套的实验指导，强调实践环节，编写了5个实验的实验指导：信道容量的迭代算法、唯一可译码判决准则、Huffman编码、LZW编码和Shannon编码，每个实验都给出了实验目的、实验要求、算法和参考程序。

第1～4章由李梅编写，第5～7章由李亦农编写，第8章由两人共同完成。

在本书的编写过程中，参阅了国内外一些经典著作(列于本书参考文献中)，同时参考了很多国外知名大学信息论基础课程的课后习题及解答，在此向有关作者表示感谢！在本书编写的过程中，得到了陈剑、李洪、徐益民、唐晓晟等同志的大力支持，在此也表示感谢！本套书在成书过程中得到了北京邮电大学出版社周明老师的热心支持，北京邮电大学田宝玉教授对教材提出了宝贵的意见，北京邮电大学信息工程学院的苏泗浠老师给出了部分习题的解答，北京邮电大学电信工程学院2000级陈敏茹同学为部分上机题编制了源程序，在此也一并表示感谢！

本书曾得到地下信息探测技术与仪器教育部重点实验室的资助。

恳请读者对书中的谬误给予指正。

作　者

2005年5月

目 录

第1章

绪　　论

《信息论基础教程》的学习要点如下：

自信息

自信息表示随机事件 x_i 发生的不确定性或发生所含有的信息量。自信息量 $I(x_i)$ 定义为

$$I(x_i) = -\log_2 p(x_i) = \log_2 \frac{1}{p(x_i)}$$

式中，$p(x_i)$为该事件发生的概率。

互信息

互信息表示已知事件 y_j 后所消除的关于事件 x_i 的不确定性，它等于事件 x_i 本身的不确定性 $I(x_i)$减去已知事件 y_j 后对 x_i 仍然存在的不确定性 $I(x_i|y_j)$。互信息 $I(x_i;y_j)$ 定义为

$$I(x_i;y_j) = I(x_i) - I(x_i|y_j) = \log_2 \frac{p(x_i|y_j)}{p(x_i)}$$

平均自信息

平均自信息表示整个信源（用随机变量 X 表示）的平均不确定性，它等于随机变量 X 的每一个可能取值的自信息 $I(x_i)$的统计平均值。平均自信息量 $H(X)$定义为

$$H(X)=E[I(x_i)]=-\sum_{i=1}^{q}p(x_i)\log_2 p(x_i)$$

式中，q 为随机变量 X 的所有可能取值的个数。

离散信源的最大熵

离散信源中各消息等概率出现时熵最大，也称为最大离散熵定理：

$$H(p_1,p_2,\cdots,p_n)\leqslant H\left(\frac{1}{n},\frac{1}{n},\cdots,\frac{1}{n}\right)=\log_2 n$$

式中，n 是随机变量 X 的所有可能取值的个数。

联合熵

联合熵表示二维随机变量 XY 的平均不确定性，它等于联合自信息的统计平均值。联合熵 $H(XY)$ 定义为

$$H(XY)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)I(x_iy_j)=-\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log_2 p(x_iy_j)$$

式中，n 和 m 分别为随机变量 X 和 Y 的所有可能取值的个数。

条件熵

条件熵表示已知随机变量 X 后，对随机变量 Y 仍然存在的平均不确定性。条件熵 $H(Y|X)$ 定义为

$$H(Y|X)=\sum_{i=1}^{n}p(x_i)H(Y|x_i)=-\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log_2 p(y_j|x_i)$$

式中，$H(Y|x_i)$ 表示已知 $X=x_i$ 的条件下，对随机变量 Y 的平均不确定性。

各类熵之间的关系：

$$H(XY)=H(X)+H(Y|X)=H(Y)+H(X|Y)\leqslant H(X)+H(Y)$$

当 X,Y 统计独立时，有

$$H(XY)=H(X)+H(Y)$$

平均互信息

平均互信息表示收到一个符号集（用随机变量 Y 表示）后所消除的关于另一个符号集（用随机变量 X 表示）的不确定性，也就是从 Y 所获得的关于 X 的平均信息量。平均互信息 $I(X;Y)$ 定义为

$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)I(x_i;y_j)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_iy_j)\log_2\frac{p(x_i|y_j)}{p(x_i)}$$

式中，n 和 m 分别为随机变量 X 和 Y 的所有可能取值的个数。

平均互信息和各类熵的关系：

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

当 X,Y 统计独立时，有

$$I(X;Y)=0$$

数据处理定理

如果随机变量 X,Y,Z 构成一个马尔可夫链，则有以下关系成立：

$$I(X;Z)\leqslant I(X;Y),\qquad I(X;Z)\leqslant I(Y;Z)$$

等号成立的条件是对于任意的 x,y,z，有 $p(x|yz)=p(x|z)$ 和 $p(z|xy)=p(z|x)$。

数据处理定理中不等式 $I(X;Z)\leqslant I(X;Y)$ 表明从 Z 所得到的关于 X 的信息量小于等于从 Y 所得到的关于 X 的信息量。如果把 $Y\rightarrow Z$ 看作数据处理系统，那么通过数据处理后，虽然可以满足我们的某种具体要求，但是从信息量来看处理后会损失一部分信息，最多保持原来获得的信息，也就是说，对接收到的数据 Y 进行处理后，决不会减少关于 X 的不确定性。

熵率

熵率表示离散多符号信源的平均不确定性，它是信源输出的符号序列中平均每个符号所携带的信息量。

如果当 $N\rightarrow\infty$ 时极限 $\lim\limits_{N\rightarrow\infty}H_N(\boldsymbol{X})$ 存在，则称之为熵率，或称为极限熵。熵率定义为

$$H_\infty=\lim_{N\rightarrow\infty}H_N(\boldsymbol{X})$$

式中，$H_N(\boldsymbol{X})$ 称为平均符号熵，它表示随机变量序列中，对前 N 个随机变量的联合熵的平均：

$$H_N(\boldsymbol{X})=\frac{1}{N}H(X_1X_2\cdots X_N)$$

离散平稳无记忆信源的熵率

多符号信源中最简单的是离散平稳无记忆信源，它的熵率为

$$H_\infty=H(X)$$

***m* 阶马尔可夫信源的熵率**

离散平稳有记忆信源中比较简单的是记忆长度有限的信源，即马尔可夫信源。如果信源在某时刻发出的符号仅与在此之前发出的 m 个符号有关，则称为 m 阶马尔可夫信源，它的熵率为

$$H_\infty=H(X_{m+1}|X_1X_2\cdots X_m)$$

$H(X_{m+1}|X_1X_2\cdots X_m)$ 常记为 H_{m+1}，它表示已知前面 m 个符号的条件下，输出下一个符号的平均不确定性。

对于离散平稳马尔可夫信源，通常将上述符号的不确定性问题转化为齐次遍历的马

尔可夫链的状态转移问题：

$$\begin{aligned}H_{m+1} &= H(X_{m+1} \mid X_1 X_2 \cdots X_m) \\ &= E[p(x_{i_{m+1}} \mid x_{i_1} x_{i_2} \cdots x_{i_m})] \\ &= E[p(x_{i_{m+1}} \mid s_i)] \\ &= -\sum_{i=1}^{q^m} \sum_{i_{m+1}=1}^{q} p(s_i) p(x_{i_{m+1}} \mid s_i) \log_2 p(x_{i_{m+1}} \mid s_i) \\ &= \sum_i p(s_i) H(X \mid s_i) \\ &= -\sum_i \sum_i p(s_i) p(s_j \mid s_i) \log_2 p(s_j \mid s_i)\end{aligned}$$

其中，$p(s_i)$是马尔可夫链的状态的平稳分布，$H(X|s_i)$表示信源处于某一状态 s_i 时发出下一个符号的平均不确定性，$p(s_j|s_i)$是状态的一步转移概率。

信源剩余度

信源剩余度定义为

$$\gamma = 1 - \eta = 1 - \frac{H_\infty}{H_0} = 1 - \frac{H_\infty}{\log_2 q}$$

其中，η 为熵的相对率，H_0 为离散信源的最大熵，$H_0 - H_\infty$越大，信源的剩余度越大。

连续信源的微分熵

$$h(X) = -\int_{\mathbf{R}} p(x) \log_2 p(x) \mathrm{d}x$$

连续信源的最大熵

对于输出信号幅度受限的连续信源，当满足均匀分布时达到最大熵；对于平均功率受限的连续随机变量，当服从高斯分布时具有最大熵。

熵功率

设某连续信源的微分熵为 $h(X)$，则将与它具有相同熵的高斯信源的平均功率 $\overline{P}$ 定义为熵功率，即

$$\overline{P} = \frac{1}{2\pi e} e^{2h(X)}$$

假定该连续信源的平均功率为 P，则 $\overline{P} \leqslant P$。熵功率和信源的平均功率相差越大，说明信源的剩余越大。信源平均功率和熵功率之差$(P - \overline{P})$称为连续信源的剩余度。

信道容量

对于给定的信道，即信道转移概率 $p(y_j|x_i)$固定后，$I(X;Y)$是 $p(x_i)$的上凸函数。

因此对于给定的信道，总存在一种信源(某种输入概率分布)，使信道平均传输一个符号接收端获得的信息量最大，也就是说对于每个固定信道都有一个最大的信息传输率，这个最大的信息传输率即信道容量，而相应的输入概率分布称为最佳输入分布。信道容量定义为平均互信息对于输入概率分布的最大值：

$$C=\max_{p(x)}\{I(X;Y)\}$$

具有扩展性能的无损信道的信道容量

$$C=\max_{p(x)}\{I(X;Y)\}=\max_{p(x)}H(X)=\log_2 r$$

具有归并性能的无噪信道的信道容量

$$C=\max_{p(x)}\{I(X;Y)\}=\max_{p(x)}H(Y)=\log_2 s$$

具有一一对应关系的无噪无损信道的信道容量

$$C=\max_{p(x)}\{I(X;Y)\}=\max_{p(x)}H(Y)=\max_{p(x)}H(X)=\log_2 s=\log_2 r$$

对称信道的信道容量

$$C=\log_2 s-H(p'_1,p'_2,\cdots,p'_s)$$

式中 $p'_1,p'_2,\cdots,p'_s$ 为信道矩阵中的任一行元素，当输入为等概分布时，达到信道容量。

准对称信道的信道容量

$$C=\log_2 r-\sum_{k=1}^{n}N_k\log_2 M_k-H(p'_1,p'_2,\cdots,p'_s)$$

式中，假定准对称信道分成 n 个对称的子信道，N_k 是第 k 个子矩阵中行元素之和，M_k 是第 k 个子矩阵中列元素之和，当输入为等概分布时，达到信道容量。

信道容量定理

设有一般离散信道，它有 r 个输入符号，s 个输出符号。当且仅当存在常数 C 使下式成立时，$I(X;Y)$达到信道容量。

$$\begin{cases}I(x_i;Y)=C, & p(x_i)\neq 0\\ I(x_i;Y)\leqslant C, & p(x_i)=0\end{cases}$$

式中 $I(x_i;Y)=\sum_j p(y_j|x_i)\log_2\frac{p(y_j|x_i)}{p(y_j)}$，它表示信道输入 x_i 时，所给出关于输出 Y 的信息量。常数 C 即为所求的信道容量。

离散平稳无记忆信道的 N 次扩展信道的信道容量

$$C^N=NC$$

独立并联信道的信道容量

当 N 个独立并联信道的信道容量都相同时，有

$$C_{并}=NC$$

级联信道的信道容量

级联信道的总的信道矩阵等于所级联信道的信道矩阵的乘积。求得级联信道的总的信道矩阵后，级联信道的信道容量就可以用求离散单符号信道的信道容量的方法计算。

波形信道的信道容量

$$C_t = B\log_2\left(1+\frac{\sigma_X^2}{N_0 B}\right)$$

这就是著名的香农公式，它适用于加性高斯白噪声信道。式中 B 为信道的带宽，N_0 为高斯白噪声的单边功率谱密度。σ_X^2 是输入随机变量 X 的方差，当随机变量 X 的均值为 0 时即为其平均功率。只有当输入信号为功率受限的高斯白噪声信号时，才能达到该信道容量。

码的分类

$$\text{码}\begin{cases}\text{非分组码(树码)}\\ \text{分组码(块码)}\begin{cases}\text{奇异码}\\ \text{非奇异码}\begin{cases}\text{非唯一可译码}\\ \text{唯一可译码}\begin{cases}\text{即时码}\\ \text{非即时码}\end{cases}\end{cases}\end{cases}\end{cases}$$

无失真定长信源编码定理

离散无记忆信源的熵为 $H(S)$，若对信源长为 N 的序列进行定长编码，码符号集中有 r 个码符号，码长为 l，则对于任意小的正数 ε，只要满足 $\frac{l}{N}\geqslant\frac{H(S)+\varepsilon}{\log_2 r}$，则当 N 足够大时，可实现几乎无失真编码，即译码错误概率为任意小。反之，如果 $\frac{l}{N}\leqslant\frac{H(S)-2\varepsilon}{\log_2 r}$，则不可能实现几乎无失真编码，当 N 足够大时，译码错误概率为 1。

Kraft 不等式(McMillan 不等式)

设信源符号集为 $S=\{s_1,s_2,\cdots,s_q\}$，码符号集为 $X=\{x_1,x_2,\cdots,x_r\}$，对信源进行编码，得到的码为 $C=\{w_1,w_2,\cdots,w_q\}$，码长分别为 $l_1,l_2,\cdots,l_q$，即时码(唯一可译码)存在的充要条件是 $\sum\limits_{i=1}^{q} r^{-l_i}\leqslant 1$，这称为 Kraft 不等式(McMillan 不等式)。

无失真变长信源编码定理(香农第一定理)

设离散无记忆信源 S 的信源熵 $H(S)$，它的 N 次扩展信源 $S^N=\{\boldsymbol{s}_1,\boldsymbol{s}_2,\cdots,\boldsymbol{s}_{q^N}\}$，其熵 $H(S^N)$，用码符号 $X=\{x_1,x_2,\cdots,x_r\}$ 对信源 S^N 进行编码，总可以找到一种唯一可译码使单个信源符号所需的平均码长满足：

$$\frac{H(S)}{\log_2 r}\leqslant\frac{\bar{L}_N}{N}<\frac{H(S)}{\log_2 r}+\frac{1}{N}$$

当 $N\to\infty$ 时，$\dfrac{\overline{L}_N}{N}=H_r(S)$，式中 $\overline{L}_N=\sum_{i=1}^{q^N}p(s_i)\lambda_i$，$\lambda_i$ 是扩展信源的信源符号 s_i 所对应的码字长度。

译码规则

设信道的输入符号集 $X=\{x_i,i=1,2,\cdots,r\}$，输出符号集 $Y=\{y_j,j=1,2,\cdots,s\}$，若对每一个输出符号 y_j 都有一个确定的函数 $F(y_j)$，使 y_j 对应于唯一的一个输入符号 x_i，则称这样的函数为译码规则，记为

$$F(y_j)=x_i \qquad i=1,2,\cdots,r;\quad j=1,2,\cdots,s$$

对于有 r 个输入、s 个输出的信道而言，输出 y_j 可以对应 r 个输入中的任何一个，所以译码规则共有 r^s 种。

错误概率

在确定译码规则 $F(y_j)=x_i$ 后，若信道输出端接收到符号 y_j，则一定译成 x_i，如果发送端发送的确实就是 x_i，就是正确译码；反之，若发送端发送的不是 x_i 就认为是错误译码。于是收到符号 y_j 条件下，译码正确概率为

$$p\,[F(y_j)\,|\,y_j]=p(x_i\,|\,y_j)$$

而译码错误概率 $p(e\,|\,y_j)=1-p(x_i\,|\,y_j)=1-p[F(y_j)\,|\,y_j]$，其中 e 表示除了 $F(y_j)=x_i$ 以外的所有符号的集合。

平均错误概率

译码后的平均错误概率 P_{E} 是译码错误概率 $p(e\,|\,y_j)$ 对 Y 的统计平均值，即

$$P_{\mathrm{E}}=E[p(e\,|\,y_j)]=\sum_{j=1}^{s}p(y_j)p(e\,|\,y_j)$$

它表示平均每接收到一个符号后的译码错误概率。

最大后验概率译码规则

选择译码函数 $F(y_j)=x^*$，使之满足条件：

$$p(x^*\,|\,y_j)\geqslant p(x_i\,|\,y_j) \qquad x^*\in X$$

则称为最大后验概率译码规则，又称为最小错误概率准则。对于每一个输出符号 y_j，$j=1,2,\cdots,s$ 均译成与之具有最大后验概率的那个输入符号 x^*，则信道译码的平均错误概率最小。

极大似然译码规则

选择译码函数 $F(y_j)=x^*$，使之满足条件：

$$p(y_j\,|\,x^*)\geqslant p(y_j\,|\,x_i) \qquad x^*\in X$$

称为极大似然译码规则。

当输入符号等概时,最大后验概率译码规则和极大似然译码规则是等价的。

费诺不等式

信道疑义度 $H(X|Y)$ 与平均错误概率 P_E 满足以下关系:

$$H(X|Y) \leqslant H(P_E) + P_E \log_2(r-1)$$

费诺不等式表明接收到 Y 后关于 X 的平均不确定性可以分为两部分:第一部分 $H(P_E)$是指接收到 Y 后是否产生错误的不确定性;第二部分 $P_E\log_2(r-1)$是已知错误 P_E 发生后,判断是哪个输入符号造成错误的最大不确定性,是$(r-1)$个符号的最大可能不确定性与 P_E 的乘积。

汉明距离

长度相同的两个符号序列(码字)$\boldsymbol{x}_i$ 与 $\boldsymbol{y}_j$ 之间的距离是指序列 $\boldsymbol{x}_i$ 和 $\boldsymbol{y}_j$ 对应位置上码元符号不同的位置的个数,称为汉明距离:

$$D(\boldsymbol{x}_i, \boldsymbol{y}_j) = \sum_{k=1}^{n} x_{i_k} \oplus y_{j_k}$$

式中 n 为符号序列的长度。

码的最小距离

码 C 中,任意两个码字的汉明距离的最小值称为该码 C 的最小距离,即

$$D_{\min} = \min\{D(\boldsymbol{w}_i, \boldsymbol{w}_j)\} \quad \boldsymbol{w}_i, \boldsymbol{w}_j \in C \text{ 且 } \boldsymbol{w}_i \neq \boldsymbol{w}_j$$

编码选择码字时,要使码的最小距离越大越好。译码时,则要将接收序列译成与其距离最小的码字,这样得到的平均错误概率最小。

有噪信道编码定理(香农第二定理)

设有一个离散无记忆平稳信道,其信道容量为 C。当待传输的信息率 $R<C$ 时,只要码长 n 足够长,则总存在一种编码可以使译码错误概率任意小。若 $R>C$,则无论 n 取多大,也找不到一种编码,使译码错误概率 P_E 任意小。

失真函数

设离散无记忆信源 X,经过信道传输后信道输出 Y,对于每一对(x_i, y_j),指定一个非负的函数 $d(x_i, y_j) \geqslant 0 (i=1,2,\cdots,r; j=1,2,\cdots,s)$,称 $d(x_i, y_j)$为单个符号的失真函数或失真度。失真函数表示信源发出一个符号 x_i,而在接收端再现为 y_j 所引起的误差或失真的大小。失真函数通常排列成矩阵形式。

平均失真度

信源的平均失真度表示某个信源通过某个信道传输后失真的大小。信源的平均失真度定义为

$$\bar{D} = E[d(x_i, y_j)] = \sum_{i=1}^{r}\sum_{j=1}^{s} p(x_i y_j) d(x_i, y_j) = \sum_{i=1}^{r}\sum_{j=1}^{s} p(x_i) p(y_j|x_i) d(x_i, y_j)$$

保真度准则

如果要求信源的平均失真度小于所允许的失真 D，即 $\bar{D}\leqslant D$，称为保真度准则。

***D* 失真许可的试验信道**

凡满足保真度准则($\bar{D}\leqslant D$)的信道称为 D 失真许可的试验信道。D 失真许可的试验信道的集合 B_D 定义为

$$B_D=\{p(y_j|x_i):\bar{D}\leqslant D\}\quad i=1,2,\cdots,r;\quad j=1,2,\cdots,s$$

信息率失真函数

对于给定的信源，即信源概率分布 $p(x_i)$固定后，$I(X;Y)$是信道转移概率 $p(y_j|x_i)$的下凸函数。因此对于给定的信源，总存在一种信道使 $I(X;Y)$达到最小。在满足保真度准则的 D 失真许可的试验信道集合 B_D 中必然有一个信道 $p(y_j|x_i)$使 $I(X;Y)$达到最小，这个最小值就是信息率失真函数，也称率失真函数。率失真函数 $R(D)$定义为

$$R(D)=\min_{p(y_j|x_i)\in B_D} I(X;Y)$$

$R(D)$是关于 D 的下凸函数，并且在定义域内是严格递减函数。其定义域的下限和上限分别为

$$D_{\min}=\sum_i p(x_i)\min_j d(x_i,y_j)$$

$$D_{\max}=\min_j\sum_i p(x_i)d(x_i,y_j)$$

限失真信源编码定理(香农第三定理)

设 $R(D)$是离散无记忆信源的信息率失真函数并且失真函数为有限值，对于任意的允许失真度 $D\geqslant 0$ 和任意的正数 $\varepsilon>0$，当信源序列长度 N 足够长时，一定存在一种编码，其码字个数 $M=\mathrm{e}^{N[R(D)+\varepsilon]}$，而编码后的平均失真度 $\bar{D}\leqslant D+\varepsilon$。反过来说，若码字个数 $M<\mathrm{e}^{NR(D)}$，则必然 $\bar{D}>D$。

第2章

信息的度量

2.1 同时掷 2 颗骰子，事件 A、B、C 分别表示：(A)仅有一个骰子是 3；(B)至少有一个骰子是 4；(C)骰子上点数的总和为偶数。试计算事件 A、B、C 发生后所提供的信息量。

解：

掷 2 颗骰子，共有 $6^2=36$ 种结果。记投掷结果为(x,y)，其中 x 和 y 是 2 颗骰子分别掷出的点数，则事件 A 对应于 10 种结果：(1,3),(2,3),(4,3),(5,3),(6,3),(3,1),(3,2),(3,4),(3,5),(3,6)；事件 B、C 分别对应于 11 种结果和 18 种结果。因此它们的概率分别为

$$p(A)=10/36$$
$$p(B)=11/36$$
$$p(C)=18/36$$

A、B、C 所提供的信息量分别为

$$I(A)=-\log_2 p(A)=1.848\,0\ \text{bit}$$

$$I(B)=-\log_2 p(B)=1.710\,5\ \text{bit}$$

$$I(C)=-\log_2 p(C)=1\ \text{bit}$$

2.2 设有 n 个球，每个球都以同样的概率 $1/N$ 落入 N 个格子($N\geqslant n$)的每一个格子中。假定：(A)某指定的 n 个格子中个各落入一个球；(B)任何 n 个格子中各落入一

球。试计算事件 A、B 发生后所提供的信息量。

解：

事件 A、B 发生的概率分别为

$$p(A)=\frac{n!}{N^n}$$

$$p(B)=\frac{C_N^n\cdot n!}{N^n}=\frac{N!}{N^n(N-n)!}$$

事件 A、B 发生提供的信息量分别为

$$I(A)=\log_2\frac{1}{p(A)}=\log_2\frac{N^n}{n!}=n\log_2 N-\log_2 n!$$

$$I(B)=\log_2\frac{1}{p(B)}=\log_2\frac{N^n(N-n)!}{N!}=n\log_2 N+\log_2(N-n)!\ -\log_2 N!$$

2.3　一信源有 4 种输出符号 $x_i, i=0,1,2,3$，且 $p(x_i)=1/4$。设信源向信宿发出 x_3，但由于传输中的干扰，接收者收到 x_3 后，认为其可信度为 0.9。于是信源再次向信宿发送该符号(x_3)，信宿无误收到。问信源在两次发送中发出的信息量各是多少？信宿在两次接收中得到的信息量又各是多少？

解：

假设信宿第一次收到的符号为 y，由于第二次发送无误收到，因此发、收信息量相等，均为

$$I(x_3|y)=-\log_2 p(x_3|y)=-\log_2 0.9=0.15\ \text{bit}$$

第一次发出的信息量为

$$I(x_3)=-\log_2 p(x_3)=-\log_2 0.25=2\ \text{bit}$$

第一次传送的信息量为两次发送信息量之差，即

$$I(x_3;y)=I(x_3)-I(x_3|y)=1.85\ \text{bit}$$

2.4　用递推性计算熵函数 $H(1/3,1/3,1/6,1/6)$ 的值。

解：

$$\begin{aligned}
H\left(\frac{1}{3},\frac{1}{3},\frac{1}{6},\frac{1}{6}\right)&=H\left(\frac{1}{3},\frac{2}{3}\right)+\frac{2}{3}\times H\left(\frac{1}{2},\frac{1}{4},\frac{1}{4}\right)\\
&=H\left(\frac{1}{3},\frac{2}{3}\right)+\frac{2}{3}\times H\left(\frac{1}{2},\frac{1}{2}\right)+\frac{2}{3}\times\frac{1}{2}\times H\left(\frac{1}{2},\frac{1}{2}\right)\\
&=H\left(\frac{1}{3},\frac{2}{3}\right)+H\left(\frac{1}{2},\frac{1}{2}\right)\\
&=1.915\ \text{比特/符号}
\end{aligned}$$

2.5 一信源有 6 种输出状态，概率分别为

$p(A)=0.5,\ p(B)=0.25,\ p(C)=0.125,\ p(D)=p(E)=0.05, p(F)=0.025$

试计算 $H(X)$。然后求消息 *ABABBA* 和 *FDDFDF* 的信息量(设信源先后发出的符号相互独立)，并将之与长度为 6 的消息序列信息量的期望值相比较。

解:

由信息熵定义，该信源输出的信息熵为

$$\begin{aligned} H(X) &= \sum_{i=1}^{6} p(x_i)\log_2\frac{1}{p(x_i)} \\ &= p(A)\log_2\frac{1}{p(A)} + p(B)\log_2\frac{1}{p(B)} + \cdots + p(F)\log_2\frac{1}{p(F)} \\ &= 0.5\log_2 2 + 0.25\log_2 4 + 0.125\log_2 8 + 2\times 0.05\log_2 20 + 0.025\log_2 40 \\ &= 1.94\ \text{比特/符号} \end{aligned}$$

消息 *ABABBA* 所含的信息量为

$$I_1 = 3I(A)+3I(B)=3[-\log_2 p(A)-\log_2 p(B)]=3(\log_2 2+\log_2 4)=9\ \text{bit}$$

消息 *FDDFDF* 所含的信息量为

$$I_2 = 3I(D)+3I(F)=3[-\log_2 p(D)-\log_2 p(F)]=3(\log_2 20+\log_2 40)=28.932\ \text{bit}$$

6 位长消息序列的信息量期望值为

$$\overline{I}=6H(X)=11.64\ \text{bit}$$

三者比较为 $I_1<\overline{I}<I_2$。

2.6 中国国家标准局所规定的二级汉字共 6 763 个。设每字使用的频度相等，求一个汉字所含的信息量。设每个汉字用一个 16×16 的二元点阵显示，试计算显示方阵所能表示的最大信息。显示方阵的利用率是多少?

解:

由于假定每个汉字的使用频度相同，它们有相同的出现概率，即

$$p(x)=\frac{1}{6\,763}$$

因此每个汉字所含的信息量为

$$I(x)=-\log_2 p(x)=-\log_2\frac{1}{6\,763}=12.7\ \text{bit}$$

每个显示方阵能显示 $2^{16\times 16}=2^{256}$种不同的状态，这 2^{256}种状态等概出现时信息熵最大，这时每种状态出现的概率为

$$p(y)=\frac{1}{2^{256}}$$

因此一个显示方阵所能显示的最大信息量是