信息论

Created By LYP

2026年

1 填空题(3*10 = 30)

1. 一个基本的数字通信系统模型通常包括信源、______________ 和信宿。
2. 在信息论中，一个事件发生的概率越小，其发生时所带来的 ______________ 就越大。
3. 当离散信源的各个符号出现概率 ______________ 时，其熵达到最大值。
4. 对于一个有 8 个等概出现符号的离散信源，其熵为 ______________ 比特/符号。
5. 线性分组码的最小汉明距离 $d_0 = 3$，则其最多可以检测 _________ 位错，最多可以纠正 _________ 位错。
6. 某编码的最小码距为 5，则它可以保证检测出 _________ 位错，或者纠正 _________ 位错。
7. 若一个码集用于前向纠错（FEC），要求能够纠正所有 1 位错误，则其最小码距至少应为 ______________。
8. ______________ 码要求任意一个码字都不是另一个码字的前缀。
9. 即时码一定是唯一可译码，但唯一可译码 _________ 是即时码。（填“是”或“不是”）
10. 给定两个二进制序列 $A = 1010110$，$B = 1001110$，它们之间的汉明距离为 ______________。
11. 码集 $C = \{00000,\ 10101,\ 01111,\ 11010\}$ 中，所有不同码字对之间汉明距离的最小值为 ______________。
12. 对于一个能纠正 1 位错误的二进制线性分组码，若信息位长度为 k，监督位长度为 r，则它们需满足的关系式（汉明界）为 ____________________________。
13. 对于 (n, k) 线性分组码，其编码效率定义为 ______________。

2 巩固题(3*10 = 30)

1. 一个完整的信息传输系统通常由__________、__________和__________组成。
2. 在信息论中，若两个事件中一个事件发生概率更小，则该事件所包含的信息量_______________。
3. 当离散信源的符号概率分布趋于 _______________ 时，其熵达到最大。
4. 若信源共有 16 个等概率符号，则该信源的熵为 _______________ 比特/符号。
5. 某线性分组码的最小汉明距离为 4，则该码最多可以检测 __________ 位错误，最多可以纠正 __________ 位错误。
6. 若某编码方案要求至少能够纠正 3 位错误，则其最小码距应不小于 _______________。
7. 某分组码的最小码距为 6，则该码可以保证检测 __________ 位错误。
8. 唯一可译码的译码结果不依赖于 _______________。
9. 即时码的一个重要特点是可以 _______________ 进行译码，而无需等待后续符号。
10. 即时码与唯一可译码的关系是：即时码 __________ 属于唯一可译码。（填"一定"或"不一定"）
11. 给定两个二进制序列
$$A = 1100110, \quad B = 1110011$$
它们之间的汉明距离为 _______________。
12. 已知码字 $a = 0010111$，$b = 0110101$，则它们之间的最小码距为 _______________。
13. 对于一个能纠正 2 位错误的线性分组码，其最小码距至少为 _______________。
14. 若某线性分组码的信息位数为 k，码长为 n，则其监督位数为 _______________。
15. 若监督位数为 r，信息位数为 k，则二进制分组码需满足的汉明界为 ____________________。

3 简答题(1*10 = 10)

根据信息论中对信源编码的分类方法，绘制信源编码的种类结构图，并在图中标明各类编码的名称及其基本关系。

要求：

- 图中应至少包括以下内容：
  - 定长编码；
  - 变长编码；
  - 唯一可译码；
  - 即时码。
- 正确反映各类编码之间的包含关系与层次结构；
- 图形清晰、层次分明，可采用树形结构或层级结构表示。

根据信息论中信息熵的相关定义，绘制信息熵关系维拉图，用以表示两个离散随机变量 X 与 Y 之间的信息关系。

要求：

- 在图中标明下列信息量：
  - $H(X)$；
  - $H(Y)$；
  - 条件熵 $H(X|Y)$；
  - 条件熵 $H(Y|X)$；
  - 平均互信息量 $I(X;Y)$。
- 正确反映各信息量之间的包含关系与相互关系；
- 维拉图结构应清晰、规范，能够直观体现信息熵的分解关系。

4 计算题(5*12 = 60)

4.1 离散信息的度量+离散信源及信源熵

题1：根据所给离散无记忆信源，计算各符号的自信息量以及该信源的信源熵。

信源如下所示：

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} a & b & c \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

在此基础上，计算符号序列 $abbca$ 的自信息量 $I(abbca)$。

题2: 根据所给离散无记忆信源，计算各符号的自信息量以及该信源的信源熵。

信源如下所示：

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ 0.4 & 0.3 & 0.2 & 0.1 \end{bmatrix}$$

并计算符号序列 $x_2x_1x_3x_2$ 的自信息量。

4.2 给定条件概率进行求解

题1： 已知离散随机变量 X 和 Y，其边缘概率分布为：

$$P(X=0)=\frac{1}{3}, \qquad P(X=1)=\frac{2}{3}$$

并且给定条件概率分布 $P(Y|X)$ 如下表所示：

$Y \backslash X$	$X=0$	$X=1$
$Y=0$	$\frac{2}{3}$	$\frac{1}{4}$
$Y=1$	$\frac{1}{3}$	$\frac{3}{4}$

试计算下列各量：

(1) 联合概率分布 $P(X,Y)$；

(2) 边缘概率分布 $P(Y)$；

(3) 熵 $H(X)$、$H(Y)$；

(4) 条件熵 $H(Y|X)$；

(5) 联合熵 $H(X,Y)$。

题2: 已知离散随机变量 X 和 Y，其边缘概率分布为：

$$P(X=0)=\frac{3}{5}, \qquad P(X=1)=\frac{2}{5}$$

并且条件概率分布 $P(Y|X)$ 如下表所示：

$Y\backslash X$	$X=0$	$X=1$
$Y=0$	$\frac{4}{5}$	$\frac{1}{2}$
$Y=1$	$\frac{1}{5}$	$\frac{1}{2}$

试求下列各量：

(1) 联合概率分布 $P(X,Y)$；

(2) 边缘概率分布 $P(Y)$；

(3) 熵 $H(X)$、$H(Y)$；

(4) 条件熵 $H(Y|X)$、$H(X|Y)$；

(5) 联合熵 $H(X,Y)$；

(6) 平均互信息量 $I(X;Y)$；

(7) 验证下列信息论关系式：

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

以及

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

4.3 给定联合概率进行求解

题1: 给定随机变量 X 和 Y 的联合概率分布如下表所示:

$Y\backslash X$	$X=0$	$X=1$
$Y=0$	$\frac{1}{4}$	$\frac{1}{4}$
$Y=1$	$\frac{1}{8}$	$\frac{3}{8}$

试计算下列各量:

(1) 边缘熵 $H(X)$、$H(Y)$;

(2) 条件熵 $H(X|Y)$、$H(Y|X)$;

(3) 联合熵 $H(X,Y)$;

(4) $H(Y)-H(Y|X)$;

(5) 平均互信息量 $I(X;Y)$。

题2: 已知两个二元随机变量 X 和 Y，其联合概率分布如下表所示：

$Y\backslash X$	$X=0$	$X=1$
$Y=0$	$\frac{1}{5}$	$\frac{2}{5}$
$Y=1$	$\frac{1}{5}$	$\frac{1}{5}$

已知条件概率公式为：

$$P(X|Y)=\frac{P(X,Y)}{P(Y)}$$

试求下列各量：

(1) 边缘熵 $H(X)$ 以及联合熵 $H(X,Y)$；

(2) 条件熵 $H(X|Y)$；

(3) 平均互信息量 $I(X;Y)$。

4.4 信源编码–霍夫曼编码

题1: 对给定离散无记忆信源进行二进制霍夫曼编码。

信源如下所示:

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} a & b & c & d & e \\ 0.30 & 0.25 & 0.20 & 0.15 & 0.10 \end{bmatrix}$$

要求:

(1) 构造该信源的二进制霍夫曼编码;

(2) 画出对应的霍夫曼编码树;

(3) 计算平均码长 L;

(4) 计算信源熵 $H(X)$;

(5) 计算编码效率 $\eta = \frac{H(X)}{L}$ 并进行分析。

题2: 给出下面离散无记忆信源的二进制霍夫曼编码。

信源如下所示:

$$\begin{bmatrix} U \\ P \end{bmatrix} = \begin{bmatrix} u_1 & u_2 & u_3 & u_4 & u_5 & u_6 \\ 0.35 & 0.20 & 0.15 & 0.10 & 0.10 & 0.10 \end{bmatrix}$$

要求:

(1) 按霍夫曼编码算法, 手写给出每一步的合并过程;

(2) 写出最终得到的霍夫曼编码结果;

(3) 画出对应的霍夫曼编码树;

(4) 计算平均码长 L 并与信源熵 $H(U)$ 比较;

(5) 计算编码效率 $\eta = \frac{H(U)}{L}$ 并分析编码性能。

4.5 汉明码–信道编码

题1： 已知 $(7,4)$ 汉明码的监督矩阵为标准形式：

$$H=\begin{bmatrix}1&0&1&1&1&0&0\\1&1&0&1&0&1&0\\0&1&1&1&0&0&1\end{bmatrix}=\left[\,P^T\mid I_3\,\right]$$

试完成下列各题：

(1) 写出对应的生成矩阵 $G=[I_4\mid P]$；

(2) 当信息序列 1101 1010 1001 送入编码器时，写出编码器输出的汉明码序列；

(3) 若接收到的码字为 111010101010，请给出译码过程，并写出译码结果。

题2: 已知另一个 $(7,4)$ 汉明码的监督矩阵为标准形式:

$$H = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \left[\, P^T \mid I_3 \,\right]$$

试完成下列各题:

(1) 写出该汉明码对应的生成矩阵 $G = [I_4 \mid P]$;

(2) 当信息序列 1011 0010 送入编码器时，写出编码器输出的汉明码序列;

(3) 若接收到的码字为 110101011010，请给出译码过程，并写出译码结果。