

2006 年度 情報科学 I レポート 3

学生用

学籍番号： _____

氏名： _____

下記の注意事項を守り、次ページ以降の問いに答え、レポートを完成させなさい。

提出期限： 2006 年 6 月 13 日 (火) 15:00 まで

提出場所： 理学部棟 正面玄関内に設置のレポートボックス

注意事項：

- (1) このページを印刷し、必要事項を記入の上 (学籍番号欄と氏名欄は 2 箇所あるので忘れずに記入すること)、レポートの表紙として提出すること。
- (2) ~~文章処理ソフトウェアや図形処理ソフトウェア等を駆使してレポートを作成し (問→解答→問→解答→… の順になるように記述すること)、A4 サイズの用紙に印刷して提出すること (手書きは不可)。~~
- (3) クラスメイトのレポートを参考にしたり、クラスメイトと協力してレポートを作成した場合は、教員控の協力者氏名欄にクラスメイトの氏名を記入すること。これらの場合も、自分の言葉で表現し直すこと。**コピー禁止。**
- (4) 情報科学 I について、あなたの声を聞かせてください (教員控の意見・質問欄に記入のこと)。気軽にどうぞ (成績には一切影響しません)。

出題者： 幸山 直人

出題日： 2006 年 5 月 31 日 (水)

得点：

/ 6

----- 切り取り線 -----

2006 年度 情報科学 I レポート 3

教員控

学籍番号： _____

氏名： _____

協力者氏名： _____ , _____ , _____

レポート作成に要した時間： _____ . _____ 時間

得点：

/ 6

意見・質問：

問 1 情報源

$$S = \left\{ \begin{array}{ccc} s_1, & s_2, & s_3 \\ p(s_1) = \frac{1}{2}, & p(s_2) = \frac{1}{4}, & p(s_3) = \frac{1}{4} \end{array} \right\}$$

について、以下の (1)～(2) の問いに答えなさい。ただし、 s_1, s_2, s_3 は独立に生起するものとする。

(1) 情報源 S を 2 進符号で符号化するとき、 S の平均符号長 L_1 を求めなさい。(1 点)

解答例

$$\begin{aligned} L_1 &= \sum_{i=1}^3 p(s_i) \log_2 \frac{1}{p(s_i)} \\ &= \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 + \frac{1}{4} \log_2 4 = \frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{4} \times 2 = 1.5 \end{aligned}$$

(2) 情報源 S^2 を 2 進符号で符号化するとき、 S^2 の平均符号長 L_2 を求めなさい。(1 点)

* 平均符号長の下限を求める問いではないので、補助定理 2.2 を使ってはならない。

（この問いは、生起確率が $1/r^k$ (k は正の整数) の形をしているとき、平均符号長が下限となっていることの確認である。すなわち、Huffman の符号化法 II を適用しても平均符号長が変化しないことを確認する。また、平均符号長が平均符号長の下限と一致しているため補助定理 2.2 が成り立つ具体例にもなっている。

解答例 平均符号長を求めるために、各 (s_i, s_j) の生起確率を求めると

$$\begin{array}{lll} (s_1, s_1) & \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}, & (s_1, s_2) & \frac{1}{2} \times \frac{1}{4} = \frac{1}{8}, & (s_1, s_3) & \frac{1}{2} \times \frac{1}{4} = \frac{1}{8} \\ (s_2, s_1) & \frac{1}{4} \times \frac{1}{2} = \frac{1}{8}, & (s_2, s_2) & \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}, & (s_2, s_3) & \frac{1}{4} \times \frac{1}{4} = \frac{1}{16} \\ (s_3, s_1) & \frac{1}{4} \times \frac{1}{2} = \frac{1}{8}, & (s_3, s_2) & \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}, & (s_3, s_3) & \frac{1}{4} \times \frac{1}{4} = \frac{1}{16} \end{array}$$

となる。従って

$$\begin{aligned} L_2 &= \sum_{i=1}^{3^2} p(\sigma_i) \log_2 \frac{1}{p(\sigma_i)} \\ &= 1 \times \left(\frac{1}{4} \log_2 4 \right) + 4 \times \left(\frac{1}{8} \log_2 8 \right) + 4 \times \left(\frac{1}{16} \log_2 16 \right) \\ &= 1 \times \frac{1}{4} \times 2 + 4 \times \frac{1}{8} \times 3 + 4 \times \frac{1}{16} \times 4 = 3 \end{aligned}$$

* 元の符号の 1 記号あたりの平均符号長に直すと $3/2 = 1.5$ となる。

評価基準 解答例に準じた解答であれば各 1 点。

問 2 テキスト 15 ページの**補助定理 2.2** の証明を完了しなさい。テキストに記した $n = 2$ の場合を参考にするとよい。(2 点)

解答例

$$\begin{aligned} H(S^n) &= \sum_{i=1}^{M^n} p(\sigma_i) \log_r \frac{1}{p(\sigma_i)} \\ &= \sum_{i_1=1}^M \sum_{i_2=1}^M \cdots \sum_{i_n=1}^M p(s_{i_1}, s_{i_2}, \dots, s_{i_n}) \log_r \frac{1}{p(s_{i_1}, s_{i_2}, \dots, s_{i_n})} \end{aligned}$$

となる。ここで、記号が独立に生起することから $p(s_{i_1}, s_{i_2}, \dots, s_{i_n}) = p(s_{i_1})p(s_{i_2}) \cdots p(s_{i_n})$ を代入すると

$$\begin{aligned} (\text{与式}) &= \sum_{i_1=1}^M \sum_{i_2=1}^M \cdots \sum_{i_n=1}^M p(s_{i_1}, s_{i_2}, \dots, s_{i_n}) \log_r \frac{1}{p(s_{i_1})p(s_{i_2}) \cdots p(s_{i_n})} \\ &= \sum_{i_1=1}^M \sum_{i_2=1}^M \cdots \sum_{i_n=1}^M p(s_{i_1}, s_{i_2}, \dots, s_{i_n}) \log_r \frac{1}{p(s_{i_1})} \\ &\quad + \sum_{i_1=1}^M \sum_{i_2=1}^M \cdots \sum_{i_n=1}^M p(s_{i_1}, s_{i_2}, \dots, s_{i_n}) \log_r \frac{1}{p(s_{i_2})} \\ &\quad \vdots \\ &\quad + \sum_{i_1=1}^M \sum_{i_2=1}^M \cdots \sum_{i_n=1}^M p(s_{i_1}, s_{i_2}, \dots, s_{i_n}) \log_r \frac{1}{p(s_{i_n})} \end{aligned}$$

となる。また、確率の性質から第 j 項 ($j = 1, 2, \dots, n$) に対して

$$p(s_{i_j}) = \sum_{i_1=1}^M \sum_{i_2=1}^M \cdots \sum_{i_{j-1}=1}^M \sum_{i_{j+1}=1}^M \cdots \sum_{i_n=1}^M p(s_{i_1}, s_{i_2}, \dots, s_{i_n})$$

が成り立つので

$$\begin{aligned} (\text{与式}) &= \sum_{i_1=1}^M \left(\sum_{i_2=1}^M \sum_{i_3=1}^M \cdots \sum_{i_n=1}^M p(s_{i_1}, s_{i_2}, \dots, s_{i_n}) \right) \log_r \frac{1}{p(s_{i_1})} \\ &\quad + \sum_{i_2=1}^M \left(\sum_{i_1=1}^M \sum_{i_3=1}^M \cdots \sum_{i_n=1}^M p(s_{i_1}, s_{i_2}, \dots, s_{i_n}) \right) \log_r \frac{1}{p(s_{i_2})} \\ &\quad \vdots \\ &\quad + \sum_{i_n=1}^M \left(\sum_{i_1=1}^M \sum_{i_2=1}^M \cdots \sum_{i_{n-1}=1}^M p(s_{i_1}, s_{i_2}, \dots, s_{i_n}) \right) \log_r \frac{1}{p(s_{i_n})} \\ &= \sum_{i_1=1}^M p(s_{i_1}) \log_r \frac{1}{p(s_{i_1})} + \sum_{i_2=1}^M p(s_{i_2}) \log_r \frac{1}{p(s_{i_2})} + \cdots + \sum_{i_n=1}^M p(s_{i_n}) \log_r \frac{1}{p(s_{i_n})} \end{aligned}$$

となる。なお、与式の各項はそれぞれ

$$\sum_{i=1}^M p(s_i) \log_r \frac{1}{p(s_i)}$$

の形をしており、 $H(S)$ に等しい。ゆえに、

$$H(S^n) = nH(S)$$

が成り立つ。 ■

評価基準 解答例に準じた解答であれば 2 点。

問 3 関数 $x \log_2 x$ ($x > 0$) について極限

$$\lim_{x \rightarrow +0} x \log_2 x$$

を求めなさい。(1 点)

解答例 $t = \frac{1}{x}$ とおき、変数変換を行なうと

$$\begin{aligned} (\text{与式}) &= \lim_{t \rightarrow +\infty} \frac{1}{t} \log_2 \frac{1}{t} \\ &= (-\log_e 2) \times \lim_{t \rightarrow +\infty} \frac{\log_e t}{t}. \end{aligned}$$

$\frac{+\infty}{+\infty}$ の不定形であるから、ロピタルの定理を利用して

$$\begin{aligned} (\text{与式}) &= (-\log_e 2) \times \lim_{t \rightarrow +\infty} \frac{(\log_e t)'}{t'} \\ &= (-\log_e 2) \times \lim_{t \rightarrow +\infty} \frac{1}{t} \\ &= 0. \end{aligned}$$

評価基準 解答例に準じた解答であれば 1 点。解答例は最も簡単な求め方である。

問 4 サイコロの出目のエントロピーを小数第 2 位まで求めなさい。(1 点)

解答例 サイコロの出目 a_i ($i = 1, 2, \dots, 6$) は等確率 $p(a_i) = \frac{1}{6}$ で起こるから、そのエントロピーは

$$H(A) = - \sum_{i=1}^6 \frac{1}{6} \log_2 \frac{1}{6} = 6 \times \frac{1}{6} \times \log_2 6 = 2.584 \dots \quad (\text{ビット})$$

となる。ゆえに、サイコロの出目のエントロピーは 2.58 ビットである (小数第 3 位を四捨五入)。

評価基準 解答例に準じた解答であれば 1 点。