後期中間試験解答用紙 (5E 情報理論)

電気工学科　　学籍番号　　　　氏名

2007 年 12 月 10 日

1　サンプリングに関する問い

[問 1] $\boxed{\textit{10 点}}$
標本化定理により，サンプリング周波数の下限は 200[KHz] である．

[問 2] $\boxed{\textit{10 点}}$
(A) の場所に挿入する．カットオフ周波数は，サンプリング周波数の半分; ナイキスト周波数，22.05 [KHz] よりも低くする必要がある．

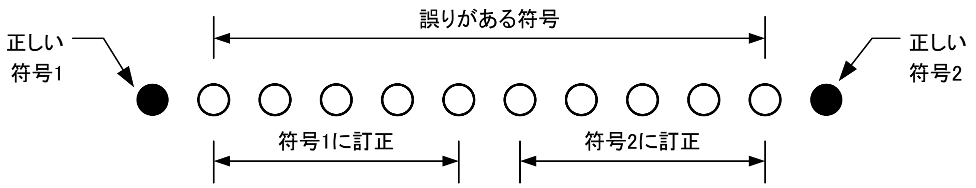
[問 3] $\boxed{\textit{5 点}}$
ローパスフィルターが無いと，22.05[KHz] の以上の信号はエイリアジングが生じ，低い周波数で再生されてしまう．

2　デジタル符号化

[問 1] $\boxed{\textit{5 点}}$
例えば，次のようにする．圧縮ソフトウェアーで圧縮し，そのデータのビット数を数える．もし，元のデータよりも大きければ，データの先頭に'0' を追加する．元のデータよりも小さければ，先頭に'1' を追加する．圧縮したデータを解凍—元に戻す—ときに，先頭のビットを見て，引き続く処理を行う．先頭のビットが'0' ならば，その先頭のビットを取り除く．一方，'1' ならば解凍の処理を行う．このようにすれば，最悪でも 1 ビットの増加で済む．

[問 2] $\boxed{\textit{10 点}}$
二つのデータで異なる符号の数は，8 である．従って，ハミング距離は，8 となる．

[問 3] $\boxed{\textit{10 点}}$
符号に t 個の誤りがある場合，正しい符号のハミング距離が $2t+1$ 以上であれば，訂正ができる．図のように，符号に誤りが生じてもハミング距離の近い正しい符号に訂正すれば良い．



以上より，正しい符号のハミング距離が $2t+1$ となる冗長性があれば，t 個のビットに誤りが生じても訂正ができることが分かる．

3 情報量

[問 1] (ア) $\boxed{\textit{10 点}}$

ぴょん太を観測して得られる平均情報量 (エントロピー)H_t は，その定義より，

$$H_t = -\sum_i p_i \log_2 p_i = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = -\log_2 2^{-1} = \log_2 2 = 1 \quad [\mathrm{bit}]$$

となる．

(イ) $\boxed{\textit{5 点}}$

前問の (ア) でぴょん太の情報量が分かったので，ぴょん吉を観測して得られる平均情報量 H_k は，

$$\begin{aligned} H_k &= -\sum_i p_i \log_2 p_i \\ &= -\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4} = -\frac{3}{4}(\log_2 3 - \log_2 4) - \frac{1}{4}\log_2 2^{-2} = 0.75\times(2-1.585)+0.5 \\ &= 0.811 \quad [\mathrm{bit}] \end{aligned}$$

となる．二匹を観測した場合の情報量は，ぴょん太とぴょん吉のそれぞれの合計となる．従って，情報量は 1.811 [bit] となる．

(ウ) $\boxed{\textit{5 点}}$

確率 p の事象が生じたときの情報量は，$\log_2 p$ である．また，情報量は加算ができるので，従って，

$$H = -\log_2\frac{1}{2} - \log_2\frac{1}{4} = 1+2 = 3\,[\mathrm{bit}]$$

となる．

別解　ぴょん太がトンボを食べる確率は 1/2，ぴょん吉がトンボを食べない確率が 1/4 である．この 2 つの事象が同時に起きる確率は 1/8 となる．したがって，同時に生じた場合の情報量は，$-\log_2(1/8) = 3\,[\mathrm{bit}]$ となる．

[問 2] (ア) $\boxed{\textit{10 点}}$

確率の合計は 1 となるので，D の出現確率 p_D は，次の通り．

$$p_D = 1 - \frac{1}{2} - \frac{1}{4} - \frac{1}{8} = \frac{1}{8}$$

(イ) $\boxed{\textit{10 点}}$

$$H = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{8}\log_2\frac{1}{8} - \frac{1}{8}\log_2\frac{1}{8} = \frac{1}{2}+\frac{2}{4}+\frac{3}{8}+\frac{3}{8} = \frac{7}{4} = 1.75 \quad [\mathrm{bit}]$$

[問 3] $\boxed{\textit{5 点}}$

1 が発生する確率を p とすると，0 が発生する確率は $1-p$ となる．この場合の平均情報量 H は，$H = -p\log_2 p - (1-p)\log_2(1-p)$ となる．これは，$0 \leqq p \leqq 1$ の間に極大値があることは，それぞれの項の関数の形より自明である．極大値は，p で微分してゼロを探せばよい．

$$\begin{aligned} \frac{\mathrm{d}H}{\mathrm{d}p} &= -\frac{1}{\log_2}\left[\log_e p + 1 - \log_e(1-p) - 1\right] \\ &= -\frac{1}{\log_2}\left[\log_e p - \log_e(1-p)\right] \end{aligned}$$

これが，ゼロに等しいので，

$$0 = \log_e p - \log_e(1-p) = \log_e\left(\frac{p}{1-p}\right)$$

となる．したがって，

$$\frac{p}{1-p} = 1$$

なので，$p = 1/2$ となる．以上より，平均情報量が最大になるのは'0' および'1' が生じる確率が 1/2 の時であることが分かる．

[問 4] $\boxed{\textit{5 点}}$

'0' と'1' が等確率で生じる N 桁の符号の平均情報量が N [bit] であることを示せばよい．N 桁の符号であるビットパターンが生じる確率は，いずれも等しく $p_i = 1/2^N$ となる．また，ビットパターンの組み合わせは 2^N 個ある．この場合の平均情報量 H は，次の通り．

$$H = -\sum_{i=1}^{2^N} = -\sum_{i=1}^{2^N} p_i\log_2 p_i = -\sum_{i=1}^{2^N}\frac{1}{2^N}\log_2\frac{1}{2^N} = 2^N\times\frac{N}{2^N} = N \quad [\mathrm{bit}]$$