情報理論 配布資料 #9

担当：井上 純一 (情報エレクトロニクス系棟 8-13)

平成 16 年 6 月 21 日

$\boxed{\text{演習問題 8 の解答例}}$

前回の $\boxed{\text{演習問題 7}}$ に引き続き, 確率変数が連続の場合に相互情報量はどのように書け, 通信路容量はどれくらいになるか, を確認しておくための演習問題である. 離散確率変数の場合, 通信路容量は相互情報量を入力分布について最大化したものとして定義されたが, 連続変数の場合も同じである. この点に注意しながら以下の小問を見ていくことにしょう.

(1) まず, 入力が特定の値 $X=x$ をとる場合の条件付きエントロピー (入力として $X=x$ を受け取ったときに出力 Y について残るあいまいさ) $H(Y|X=x)$ は

$$
\begin{aligned}
H(Y|X=x) &= -\int_{-\infty}^{\infty} dy P_{Y|X}(y|x)\log P_{Y|X}(y|x) \\
&= -\int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi\sigma_N^2}}\,\mathrm{e}^{-\frac{(y-x)^2}{2\sigma_N^2}} \log\left\{\frac{1}{\sqrt{2\pi\sigma_N^2}}\,\mathrm{e}^{-\frac{(y-x)^2}{2\sigma_N^2}}\right\} \\
&= -\frac{\log \mathrm{e}}{2\sigma_N^2}\int_{-\infty}^{\infty}\frac{dy}{\sqrt{2\pi\sigma_N^2}}(y-x)^2\mathrm{e}^{-\frac{(y-x)^2}{2\sigma_N^2}} - \frac{1}{2}\log(2\pi\sigma_N^2) \\
&= -\frac{\log \mathrm{e}}{2\sigma_N^2}\int_{-\infty}^{\infty}\frac{dz}{\sqrt{2\pi\sigma_N^2}}\,z^2\mathrm{e}^{-\frac{z^2}{2\sigma_N^2}} - \frac{1}{2}\log(2\pi\sigma_N^2) \qquad (1)
\end{aligned}
$$

となる. 最後の変形では簡単のため, $z \equiv y-x$ とおいた. さて, ここから先は上式右辺の第 1 項目の次のタイプの積分：

$$
\int_{-\infty}^{\infty} dz\,z^2\,\mathrm{e}^{-az^2} = -\frac{\partial}{\partial a}\left\{\int_{-\infty}^{\infty} dz\,\mathrm{e}^{-az^2}\right\} \qquad (2)
$$

を評価することが必要となるが, 前回の演習問題で示した公式：

$$
\int_{-\infty}^{\infty} dz\,\mathrm{e}^{-az^2} = \sqrt{\frac{\pi}{a}} \qquad (3)
$$

を (2) 式の $\{\cdots\}$ に代入して, a に関する微分を実行すれば

$$
\int_{-\infty}^{\infty} dz\,z^2\,\mathrm{e}^{-az^2} = \frac{\sqrt{\pi}}{2a\sqrt{a}} \qquad (4)
$$

が得られるから, (1) 式で $a=1/2\sigma_N^2$ とおくことにより, $H(Y|X=x)$ は

$$
H(Y|X=x) = \frac{1}{2}\log\mathrm{e} + \frac{1}{2}\log(2\pi\mathrm{e}\sigma_N^2) = \frac{1}{2}\log(2\pi\mathrm{e}\sigma_N^2) \qquad (5)
$$

となる．この結果は 2 元対称通信路 (BSC) の場合と同様に，入力値 $X = x$ には依らないことに注意しょう．従って，$H(Y|X = x)$ を $X = x$ に関する確率分布 $P_X(x)$ で平均した条件付きエントロピー $H(Y|X)$ は $P_X(x)$ に依らず

$$
\begin{aligned}
H(Y|X) &= \int_{-\infty}^{\infty} dx P_X(x) H(Y|X = x) \\
&= \frac{1}{2}\log(2\pi \mathrm{e}\sigma_N^2)\int_{-\infty}^{\infty} dx P_X(x) = \frac{1}{2}\log(2\pi \mathrm{e}\sigma_N^2) \qquad (6)
\end{aligned}
$$

となる．

(2) 定義に従って，$\langle y^2 \rangle$ を書き出してみると

$$
\begin{aligned}
\langle y^2 \rangle &= \int_{-\infty}^{\infty} dx P_X(x) \int_{-\infty}^{\infty} \frac{y^2 dy}{\sqrt{2\pi\sigma_N^2}} \mathrm{e}^{-\frac{(y-x)^2}{2\sigma_N^2}} \\
&= \int_{-\infty}^{\infty} dx\, x^2 P_X(x) \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi\sigma_N^2}} \mathrm{e}^{-\frac{z^2}{2\sigma_N^2}} + \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi\sigma_N^2}} \mathrm{e}^{-\frac{z^2}{2\sigma_N^2}} \int_{-\infty}^{\infty} dx P_X(x) \\
&= \langle x^2 \rangle + \sigma_N^2 \leq \sigma^2 + \sigma_N^2 \qquad (7)
\end{aligned}
$$

となる．ここで，1 行目から 2 行目への変形では $y - x = z$ と変換し，

$$
\int_{-\infty}^{\infty} dz\, z\, \mathrm{e}^{-az^2} = \left[-\frac{1}{2a} z^2\right]_{-\infty}^{\infty} = 0 \qquad (8)
$$

という事実を使い，2 行目の右辺第 2 項の積分は (4) を用いた．また，最後の不等式は確率変数 x の 2 乗平均が σ^2 であることから明らかである．よって題意の式が示された．

(3) 最後に通信路容量の評価であるが，今の場合，相互情報量 $I(X;Y)$ が

$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y|X) \\
&= H(Y) - \frac{1}{2}\log(2\pi \mathrm{e}\sigma_N^2) \qquad (9)
\end{aligned}
$$

であるから，あとは $H(Y)$ を最大化すればよいが，前回の演習問題の結果と (2) の結果を合わせると，エントロピー $H(Y)$ は確率変数 $X = x$ の $P_X(x)$ に関する 2 乗平均を用いて

$$
\begin{aligned}
H(Y) &\leq \frac{1}{2}\log(2\pi \mathrm{e}\langle y^2 \rangle) \\
&\leq \frac{1}{2}\log(2\pi \mathrm{e}(\sigma^2 + \sigma_N^2)) \qquad (10)
\end{aligned}
$$

と上から押さえられるので，結局

$$
\begin{aligned}
I(X;Y) &\leq \frac{1}{2}\log(2\pi \mathrm{e}(\sigma^2 + \sigma_N^2)) - \frac{1}{2}\log(2\pi \mathrm{e}\sigma_N^2) \\
&= \frac{1}{2}\log\left(1 + \frac{\sigma^2}{\sigma_N^2}\right) = C \qquad (11)
\end{aligned}
$$

が得られる．

演習問題 7 に関するコメント

こちらの誘導が悪かったのでしょうか．全体的にあまり出来が良くありませんでした．解答例に示したように，本配布資料の (4) 式という積分結果を使わなくてもできるのですが，この積分が出来ないために頓挫している答案が散見されました．また，前回配布の解答例の **(参考)** に示したラグランジュの未定係数法を用いて解いて下さった方が 1 名いましたが，この方法は何か別の講義で既習なのでしょうか？ それともこの解答者が自習したのでしょうか？ いずれにしても，便利な方法ではあるので，覚えておいて損は無いと思います．
今回と前回の 2 回に渡って連続変数の場合のエントロピー，相互情報量等の扱い方を見てきましたが，講義ではシャノンの通信路符号化定理をやっているわけで，講義内容と演習内容が一致していなかったこともあり，皆さんは問題に取り組みにくかったのではないかな，と，こちらとしては察するわけですが，当初の予定通り連続信号のところを講義で説明するには時間が足りず，そのためにやむを得ず「演習問題」という形でこの部分を補う方法を採った次第です．

ところで，この 2 回の演習問題で連続確率変数のエントロピーを

$$H(X) \quad = \quad -\int_{-\infty}^{\infty} dx P_X(x) \log P_X(x) \tag{12}$$

のように定義して使ったのですが，「これは既に学んだ離散変数のエントロピーに対する何らかの連続極限として導けるだろうか？」という疑問を持つのはとても健全なことです．しかし，これをやろうとすると実は厄介な問題に首を突っ込むことになるので演習問題ではあえて避けていたのですが (問題文に含めるとダラダラした，しまりのない文章になりかねない !)，ここでそれについて多少のコメントをしておきましょう．

まずは x が離散値 x_i $(i=-\infty,\cdots,-1,0,1,\cdots,\infty)$ をとるものし，$\Delta x \equiv x_{i+1}-x_i$ と定義すると，確率変数が区間 (x_i, x_{i+1}) にある値をとる確率は $P_X(x_i)\Delta x$ なので，この離散変数のエントロピーは従来通りに定義でき，ここで $\Delta x \to 0$ の極限をとることによって連続変数のエントロピーの定義とすれば

$$\begin{aligned} H(X) &\equiv -\lim_{\Delta x\to 0}\sum_{i=-\infty}^{\infty} P_X(x_i)\Delta x \log\{P_X(x_i)\Delta x\} \\ &= -\lim_{\Delta x\to 0}\sum_{i=-\infty}^{\infty} \Delta x\, P_X(x_i)\log P_X(x_i) - \lim_{\Delta x\to 0}\sum_{i=-\infty}^{\infty} \Delta x\, P_X(x_i)\log \Delta x \\ &= -\int_{-\infty}^{\infty} dx\, P_X(x)\log P_X(x) - \lim_{\Delta x\to 0}\log \Delta x \end{aligned} \tag{13}$$

となります．しかし，この結果から見て取れるように $\Delta x \to 0$ **の連続極限と取ったとたん，上式第 1 項は所望のエントロピーを与えますが，第 2 項が** $+\infty$ **に発散してしまいます**．上の極限のとり方自体に間違いは無かったので，我々はここで「この発散項を一体どうしてくれるのか？」という問題に直面するわけです．

この講義の最初に学びましたが，エントロピーとは「事象 X に関するあいまいさ」，あるいは「X が起こることを当てる難易度」でした．これを思い出すと，上で生じた問題は変数 $X=x$ を連続実数に拡張したことにより，言ってみれば「無限個の候補」を作ってしまい，その結果「無限個の候補」の中から正解一つを当てる難易度が無限大となってしまった結果である，として理解できます．こう考えるとこの結論 (発散) 自体は当たり前のような気がしてきますが，一方でエントロピーを具体的に計算する際には $-\lim_{\Delta x\to 0}\log(\Delta x)$ の値をどうするのか，という実用的な問題もあるかと思います．しかしその答えをここで言うならば「どうでもいい」ということになります．この演習問題では暗黙のうちにこの値をゼロと置いていました．

重要なのは**連続変数の場合, エントロピー自体の値には意味がなく, その差にこそ意味がある**という点です. **演習問題 7** で $H(X)$ と $P_X(x)$ が正規分布の場合の値 $\log(2\pi e\sigma^2)/2$ との「差」を評価したのはそのためです (実際, $\log(2\pi e\sigma_2)/2$ の値自体は正規分布の分散が $\sigma \le 1/\sqrt{2\pi e}$ の時には負になります). そして最大値は必ずその他の値との差をとれば正の値となりますから, エントロピーの最大値を与える分布 (ここでは正規分布) 自体は意味を持つわけです.

以上のことに注意しておけば, 残るのは技術的な問題です. さらに後から見る「標本化定理」を学べば, 連続情報 (信号) の扱い方としては大体 OK ということになるかと思います.

演習問題 9

成分が 0,1 である任意のベクトル $\boldsymbol{u} = (u_1, u_2, \cdots, u_n), \boldsymbol{v} = (v_1, v_2, \cdots, v_n)$ に対し, 和を $\boldsymbol{u} + \boldsymbol{v} = (u_1 \oplus v_1, u_2 \oplus v_2, \cdots, u_n \oplus v_n)$ で定義する. ここで, 記号 $\oplus$ は排他的論理和を表し, 交換則 $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ が成り立つ. また, ベクトルにスカラー $a \in \{0,1\}$ をかける演算を $a\boldsymbol{u} = (au_1, au_2, \cdots, au_n)$ で定義し, 2 つのスカラーの積は $00 = 01 = 10 = 0, 11 = 1$ とする. つまり, $a = 0$ のとき $a\boldsymbol{u} = \boldsymbol{0}$, $a = 1$ のとき $a\boldsymbol{u} = \boldsymbol{u}$ である. ここに $\boldsymbol{0} = (0, 0, 0, \cdots, 0)$ はゼロベクトルである. これらの演算が成り立つベクトルに対して以下の問に答えよ.

(1) 任意のベクトル $\boldsymbol{v}$ に対し, $\boldsymbol{v} + \boldsymbol{e} = \boldsymbol{v}$ となるような $\boldsymbol{e}$ は何か.

(2) 任意のベクトル $\boldsymbol{v}$ に対し, $\boldsymbol{v} + \boldsymbol{f} = \boldsymbol{0}$ となるような $\boldsymbol{f}$ は何か.

(3) 任意のベクトル $\boldsymbol{u}$ の成分中 $u_i \neq 0$ となる i の個数をハミング重みと呼び $w(\boldsymbol{u})$ で表す. このとき任意のベクトル $\boldsymbol{u}, \boldsymbol{v}$ に対し, 不等式 :

$$w(\boldsymbol{u}) + w(\boldsymbol{v}) \quad \geq \quad w(\boldsymbol{u} + \boldsymbol{v})$$

が成り立つことを示せ.

(4) パリティ検査行列 :

$$\boldsymbol{H} \quad = \quad \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

を持つ線形符号に対し, 1 ビット誤りベクトルの復号表を作成せよ. つまり, 1 ビット誤りベクトル $\boldsymbol{z} = (1,0,0,0,0,0), (0,1,0,0,0,0), (0,0,1,0,0,0), \cdots, (0,0,0,0,0,1)$ に対して, $\boldsymbol{z}$ と $\boldsymbol{s}_z \equiv \boldsymbol{z}\boldsymbol{H}^T$ の対応表を作れ. ここに $\boldsymbol{H}^T$ は $\boldsymbol{H}$ の転置である. また, この復号表から, この線形符号が 1 ビット誤り訂正符号であるか判定せよ.

(5) 送信ベクトル $\boldsymbol{u}$ が通信路を通過後, 1 ビットノイズ $\boldsymbol{z}$ が加わり, $\boldsymbol{r} = \boldsymbol{u} + \boldsymbol{z}$ として受信された. 受信ベクトル $\boldsymbol{r}$ に対し, $\boldsymbol{s}_r \equiv \boldsymbol{r}\boldsymbol{H}^T$ を計算し, (4) で作成した復号表から $\boldsymbol{s}_z = \boldsymbol{s}_r$ となる $\boldsymbol{z}$ を求めると, 送信ベクトルは $\boldsymbol{u} = \boldsymbol{r} + \boldsymbol{z}$ として復元できる. この方法を用いて受信ベクトルが $\boldsymbol{r} = (0,1,0,1,1,0)$ のとき, $\boldsymbol{u}$ 及び $\boldsymbol{z}$ を求めよ.

(6) この符号を誤り率 p の 2 元対称通信路で用いたときの誤り確率 p_z を求めよ.