情報理論 配布資料 #11

担当：井上 純一 (情報科学研究科棟 8-13)
URL：http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/

平成 17 年 7 月 4 日

目 次

演習問題 10 の解答例

いろいろなやり方がある.

$g(x)=x^3+x^2+1$ は既約多項式であるから, この解 α：$\alpha^3+\alpha^2+1=0$ を用いて, α の冪を書き出して行くと

$$\alpha^0 = 1+0\cdot\alpha+0\cdot\alpha^2 \tag{326}$$
$$\alpha^1 = 0+1\cdot\alpha+0\cdot\alpha^2 \tag{327}$$
$$\alpha^2 = 0+0\cdot\alpha+1\cdot\alpha^2 \tag{328}$$
$$\alpha^3 = 1+0\cdot\alpha+1\cdot\alpha^2 \tag{329}$$
$$\alpha^4 = 1+1\cdot\alpha+1\cdot\alpha^2 \tag{330}$$
$$\alpha^5 = 1+1\cdot\alpha+0\cdot\alpha^2 \tag{331}$$
$$\alpha^6 = 0+1\cdot\alpha+1\cdot\alpha^2 \tag{332}$$

$\alpha^7=\alpha^0$ と書けるので, この生成多項式 $g(x)$ は原始既約多項式であり, パリティ検査行列は 7×3 の行列で具体的に

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix} \tag{333}$$

となり, 符号語ベクトル $\boldsymbol{x}=(x_0,\cdots,x_6)$ に対して, パリティ検査方程式 $\boldsymbol{H}\boldsymbol{x}^{'}=\boldsymbol{0}$ は

$$x_0 = x_3+x_4+x_5 \tag{334}$$
$$x_1 = x_4+x_5+x_6 \tag{335}$$
$$x_2 = x_3+x_4+x_6 \tag{336}$$

であり, 情報ビット ($x_3x_4x_5x_6$ の 4 ビット) を決めるとパリティ検査ビット ($x_0x_1x_2$) が決まり, 最小距離は明らかに 3 であるので, $(7,4,3)$ ハミング符号である (ハミング不等式を等式で満たす).

具体的に $M=2^4=16$ 個の符号 C を書き出してみると

$$\begin{aligned} C &= \Big\{(0000000),(0110001),(1100010),(1010011),\\ &(1110100),(1000101),(0010110),(0100111),\\ &(1011000),(1101001),(0111010),(0001011),\\ &(0101100),(0001101),(1001110),(1111111)\Big\} \end{aligned} \tag{337}$$

となる. また, $H=(\alpha^0,\alpha^1,\cdots,\alpha^6)$ とおけば, パリティ検査方程式は

$$(\alpha^0,\alpha^1,\cdots,\alpha^6)\begin{pmatrix} x_0\\ x_1\\ x_2\\ \cdots\\ \cdots\\ x_6 \end{pmatrix} = 0 \tag{338}$$

つまり

$$\alpha^0x_0+\alpha^1x_1+\cdots+\alpha^6x_6 = 0 \tag{339}$$

であるから, この両辺に次々と α をかけて行くことにより, 任意の符号語ベクトル：$(x_0,x_1,\cdots,x_6)$ の巡回ベクトルである $(x_6,x_0,x_1,\cdots,x_5),\cdots$ もパリティ検査方程式を満たす符号語であることがわかる. 以上より, 生成多項式が $g(x)=x^3+x^2+1$ で与えられる $(7,4,3)$ 線形符号は巡回 (ハミング) 符号であることがわかる.

また, 上記のようにパリティ検査行列に直さなくとも, $b(x)$ を x に関する 3 次以下の多項式とすると, 生成多項式が $g(x)$ で与えられる任意の符号語は $b(x)g(x)$ で与えられることを用いれば, 各々の符号語を多項式で表現することができる. $f(x)=b(x)g(x)$ とおき, この積の多項式の係数を

$$f(x) = x_0x^6+x_1x^5+x_2x^4+x_3x^3+x_4x^2+x_5x+x_6 \tag{340}$$

として $(x_0x_1\cdots x_6)$ を符号語とすれば具体的に次の表が出来上がる.

$b(x)$	符号語の多項式表現：$f(x)=b(x)g(x)$	符号語：$(x_0x_1x_2x_3x_4x_5x_6)$
0	0	(0000000)
1	x^3+x^2+1	(0001101)
x	x^4+x^3+x	(0011010)
x^2	$x^5+x^4+x^2$	(0110100)
x^3	$x^6+x^5+x^3$	(1101000)
$1+x$	x^4+x^2+x+1	(0010111)
$1+x^2$	$x^5+x^4+x^3+1$	(0111001)
$1+x^3$	$x^6+x^5+x^2+1$	(1100101)
$x+x^2$	$x^5+x^3+x^2+x$	(0101110)
$x+x^3$	$x^6+x^5+x^4+x$	(1110010)
x^2+x^3	$x^6+x^4+x^3+x^2$	(1011100)
$1+x+x^2$	x^5+x+1	(0100011)
$x+x^2+x^3$	x^6+x^2+x	(1000110)
$1+x^2+x^3$	x^6+x^4+1	(1010001)
$1+x+x^3$	$x^6+x^5+x^4+x^3+x^2+x+1$	(1111111)
$1+x+x^2+x^3$	x^6+x^3+x+1	(1001011)

のように計 16 個の符号語が求まることになる．もちろん，これらの結果はパリティ検査方程式から求めたものと一致する．

9 連続量の情報

今まで学んできた情報理論では，個々の事象には番号をつけることができ，数え上げることができるような場合を扱ってきた．ここでは全ての情報は 0,1 で表現されたものであり，その並びで個々の通報の記号を表した．しかし，音声などを計測する場合には信号の電流値という，「実数値」に対して確率分布が定義される必要性がでてくる．ここでは，そのような連続量に対して各種の情報量がどのように定義されるのか，について詳しく見ていくことにする．

9.1 連続量のエントロピー

ここではまず，エントロピー (平均情報量，自己情報量) を連続量まで許した事象に拡張することを考える．そこで，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

$$
\begin{aligned}
H(X) &\equiv -\lim_{\Delta x\to 0}\sum_{i=-\infty}^{\infty} P_X(x_i)\Delta x\log\{P_X(x_i)\Delta x\} \\
&= -\lim_{\Delta x\to 0}\sum_{i=-\infty}^{\infty}\Delta x\,P_X(x_i)\log P_X(x_i) - \lim_{\Delta x\to 0}\sum_{i=-\infty}^{\infty}\Delta x\,P_X(x_i)\log\Delta x \\
&= -\int_{-\infty}^{\infty} dx\,P_X(x)\log P_X(x) - \lim_{\Delta x\to 0}\log\Delta x
\end{aligned}
\tag{341}
$$

となる (ここで，$\sum_{i=-\infty}^{\infty}\Delta x\,P_X(x)=1$ であることに注意)．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，第 2 項が $+\infty$ に発散してしまう**ことがわかる．この講義の最初に学んだよう

に, 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」であった. 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特定する難易度が無限大となってしまった結果である, として理解できる. こう考えるとこの結論 (発散) 自体は自然に理解できるが, 一方でエントロピーを具体的に計算する際には $-\lim_{\Delta x \to 0} \log(\Delta x)$ の値を具体的にどう処理するのか, という問題が残る. しかし, その値をいくつに置くかは本質的な問題ではなく, ケース・バイ・ケースで適当な値を選んでよい (通常はゼロとおく場合が多いように思われる) [1]. 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**という点である.

連続量のエントロピーに慣れるために次のような例題を見ておこう.

例題 11

確率変数 X が $0, 1$ のように離散的な値ではなく, 連続値をとるような確率分布 $P_X(x)$ を考えよう. この分布 $P_X(x)$ に対し, エントロピーは

$$H(X) \;=\; -\int_{-\infty}^{\infty} P_X(x) \log_2 P_X(x)\, dx \tag{342}$$

で定義される. 分布 $P_X(x)$ が 2 乗平均値 σ^2 を持つものとして以下の問いに答えよ.

(1) 分布 $P_X(x)$ が平均 0, 分散 σ^2 の正規分布に従う場合, エントロピー $H(X)$ を求めよ.

(2) $\alpha\,(>0)$ に対し, 次の不等式が成り立つことを示せ.

$$\log_2 \alpha \le (\alpha - 1)\, \log_2 \mathrm{e} \tag{343}$$

ただし, e は自然対数の底である.

(3) (1)(2) の結果を用いて

$$H(X) \;\le\; \frac{1}{2} \log_2 (2\pi \mathrm{e}\, \sigma^2) \tag{344}$$

が成り立つことを示し, 等号成立する条件を求めよ.

(※ 注) これらの問題を解く際に必要であるならば次の積分公式 :

$$\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-ax^2} \;=\; \sqrt{\frac{\pi}{a}} \tag{345}$$

を用いてもよい.

(解答例)

[1] この辺りの事情は 1 年生のときに物理学 I (力学) で学んだ「位置エネルギー」と似ている. 位置エネルギーの値自体はその基準をどこに選ぶかによって任意性があったが, いつでもその「差」には力学的な意味があり, エネルギーが保存する状況下ではこの差が運動エネルギーの増加 (減少) 分に変化する, というようなことを学んだはずである.

既に 2 値エントロピー関数, 及び, それを一般化した n 値エントロピー関数：

$$H = -\sum_{i=1}^{N} p(x_i)\log p(x_i) \tag{346}$$

を学習し, 確率変数 x_i, $(i=1,\cdots,n)$ の取りうる n 個の値 (事象) が全て等確率で現れる場合, つまり, $p(x_i)=1/n$, $(i=1,\cdots,n)$ のとき H はその最大値 $\log n$ をとることを学んだ. n 値エントロピー関数の確率変数 X は $x_1,x_2,\cdots,x_n$ のような n 個の離散値であったが, この X が連続値をとり, 分布 $P_X(x)$ に従って, その 2 乗平均値が

$$\int_{-\infty}^{\infty} dx\, x^2\, P_X(x) = \sigma^2 \tag{347}$$

で与えられる場合, この連続的確率変数 X のエントロピー：

$$H(X) = -\int_{-\infty}^{\infty} dx\, P_X(x)\log_2 P_X(x) \tag{348}$$

の最大値と, それを与える分布 $P_X(x)$ とはいったいどのような分布なのか, を問うているのがこの問題である. 問題文には既に結論：「$P_X(x)$ が平均ゼロ, 分散 σ^2 の正規分布のときにエントロピー $H(X)$ は最大値をとる」が与えられており, これを示すことがここでの課題となっている. 小問化された誘導に従い, これを見ていくことにしょう.

(1) $P_X(x)$ は平均ゼロ, 分散 σ^2 の正規分布であるということなのだから

$$P_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{x^2}{2\sigma^2}} \tag{349}$$

であり, この $P_X(x)$ に対し, エントロピー $H(X)$ は

$$\begin{aligned} H(X) &= -\int_{-\infty}^{\infty} dx\, P_X(x)\log_2\left\{\frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\right\} \\ &= -\{\log_2\mathrm{e}\}\int_{-\infty}^{\infty} dx\, P_X(x)\left\{-\frac{x^2}{2\sigma^2}\right\} + \frac{1}{2}\log_2(2\pi\sigma^2)\int_{-\infty}^{\infty} dx\, P_X(x) \\ &= \frac{\log_2\mathrm{e}}{2\sigma^2}\int_{-\infty}^{\infty} dx\, x^2\, P_X(x) + \frac{1}{2}\log_2(2\pi\sigma^2)\int_{-\infty}^{\infty} dx\, P_X(x) \\ &= \frac{1}{2}\log_2(2\pi\mathrm{e}\sigma^2) \end{aligned} \tag{350}$$

となる. ここで, 条件 (347), 及び, 確率の規格化条件：

$$\int_{-\infty}^{\infty} dx\, P_X(x) = 1 \tag{351}$$

を用いた. また, 対数の底としては 2 を使っていることに注意されたい.

(2) 対数関数の底の変換規則に従えば

$$\frac{\log_2\alpha}{\log_2\mathrm{e}} = \log\alpha \tag{352}$$

であり, $y=\log\alpha$ に $\alpha=1$ で接する直線が $y=\alpha-1$ であり, $\log\alpha$ は上に凸の関数であるから

$$\log\alpha \le \alpha-1 \tag{353}$$

が成り立つ.

(3) $H(x)$ と $H(x)$ を入力分布 $P_X(x)$ が正規分布の場合に計算した値：$H_0 = \log_2(2\pi e\sigma^2)/2$ との差を評価していこう.

$$
\begin{aligned}
H(X) - \frac{1}{2}\log_2(2\pi e\sigma^2) &= -\int_{-\infty}^{\infty} dx\, P_X(x)\log P_X(x) + \int_{-\infty}^{\infty} dx\, P_X(x)\log\left\{\frac{1}{\sqrt{2\pi\sigma^2}}\,e^{-\frac{x^2}{2\sigma^2}}\right\} \\
&= \int_{-\infty}^{\infty} dx\, P_X(x)\log\left\{\frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}P_X(x)}\right\} \qquad (354)
\end{aligned}
$$

となるが, 上式の対数の中身 $\{\cdots\}$ が α だと思って, (2) で示した不等式を使えば

$$
\begin{aligned}
H(X) - \frac{1}{2}\log_2(2\pi e\sigma^2) &\le \int_{-\infty}^{\infty} dx\, P_X(x)\left\{\frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}P_X(x)} - 1\right\} \\
&= \int_{-\infty}^{\infty} dx\, \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} - \int_{-\infty}^{\infty} dx\, P_X(x) = 1 - 1 = 0 \qquad (355)
\end{aligned}
$$

従って

$$
H(X) \le \frac{1}{2}\log_2(2\pi e\sigma^2) \qquad (356)
$$

であり, 等号が成立するのは $P_X(x)$ が平均ゼロ, 分散 σ^2 の正規分布に従うときである. 先に連続量のエントロピーはその値自体には意味が無く, その差にこそ意味があると述べたが, ここでは $H(X)$ 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 (実際, $\log(2\pi e\sigma_2)/2$ の値自体は正規分布の分散が $\sigma \le 1/\sqrt{2\pi e}$ の時には負になる). 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るから, 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ある. ここで得られた結果 (356) は, 後に連続量の通信路容量を求める際に用いるので覚えておこう.

(別解)

この問題では, 連続変数 X のエントロピー $H(X)$ を最大化するような分布 $P_X(x)$ を求めたわけだが, 「結局のところ解が正規分布であること知っていなければできないのでは ?」「直観的に解が正規分布であろう, と察しをつけることはちょっと難しい」「条件が少しでも変わったらどうなるの ?」「とにかくはじめに正解ありき, という感じで気に入らない」という人もいるかもしれない. n 値エントロピー関数のところでは, ある制約条件の下で多変数関数の最大化を行う方法として**ラグランジュの未定係数法**というのを**付録**につけたが, それを思い出せば, ここでの連続確率変数の場合にも適用することができる. この方法を使えばストレートに $P_X(x)$ を見つけることができるので, 別解としてそれを示しておく.

まず, 条件 (351)(347) の下で $H(x)$ を最大化するわけですから, 未定係数を λ, β として

$$
\begin{aligned}
F(P_X(x), \lambda, \beta) &= -\int_{-\infty}^{\infty} dx\, P_X(x)\log P_X(x) + \lambda\left\{\int_{-\infty}^{\infty} dx\, P_X(x) - 1\right\} \\
&+ \beta\left\{\int_{-\infty}^{\infty} dx\, x^2 P_X(x) - \sigma^2\right\} \qquad (357)
\end{aligned}
$$

を作れば, 我々が最大化すべきなのは関数 F ということになる. ただし, n 値エントロピー関数の場合と少し事情が違っているのは, F の中には変数として関数 $P_X(x)$ が含まれており, F は「関数の関数」— こういうのを**汎関数**と言いますが — になっている点である.

しかし, ここではあまり細かいことを気にせずに $P_X(x), \lambda, \beta$ に関して F の極値条件を書き出してみると

$$\frac{\delta F}{\delta P_X(x)} = -\int_{-\infty}^{\infty} dx \left\{\log P_X(x) + 1 + \lambda + \beta x^2\right\} = 0 \tag{358}$$

$$\frac{\partial F}{\partial \lambda} = \int_{-\infty}^{\infty} dx\, P_X(x) - 1 = 0 \tag{359}$$

$$\frac{\partial F}{\partial \beta} = \int_{-\infty}^{\infty} dx\, x^2 P_X(x) - \sigma^2 = 0 \tag{360}$$

が得られる．ここで, 関数 $P_X(x)$ での微分は他の 2 つの変数 λ, β での微分とはやや意味が違うので (「汎関数微分」), 別の微分記号 δ を用いてある．

さて, 上の 2,3 番目の式からは条件 (351)(347) が再度得られたことになる．従って, まずは問題となるのが 1 番目の式だが, これが任意の x について成り立つためには被積分関数の $\{\cdots\}$ がゼロとなるべきなので

$$P_X(x) = \mathrm{e}^{-1-\lambda-\beta x^2} \tag{361}$$

が得られる．あとは λ, β を決めるのだが, この $P_X(x)$ を (359) に代入すれば

$$\mathrm{e}^{-1-\lambda} = \left(\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\beta x^2}\right)^{-1} \tag{362}$$

となる．これを用いて (361) を書き直すと

$$P_X(x) = \frac{\mathrm{e}^{-\beta x^2}}{\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\beta x^2}} \tag{363}$$

が得られるが, これを (360) に代入すると

$$\frac{\int_{-\infty}^{\infty} dx\, x^2 \mathrm{e}^{-\beta x^2}}{\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\beta x^2}} = \sigma^2 \tag{364}$$

すなわち

$$-\frac{\partial}{\partial \beta}\left\{\log \int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\beta x^2}\right\} = \sigma^2 \tag{365}$$

が導ける．例題文中に与えた積分公式から

$$\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\beta x^2} = \sqrt{\frac{\pi}{\beta}} \tag{366}$$

なので, 結局 β が

$$\beta = \frac{1}{2\sigma^2} \tag{367}$$

と定まり, $P_X(x)$ は

$$P_X(x) = \frac{\mathrm{e}^{-\frac{x^2}{2\sigma^2}}}{\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\frac{x^2}{2\sigma^2}}} = \frac{1}{\sqrt{2\pi\sigma^2}}\, \mathrm{e}^{-\frac{x^2}{2\sigma^2}} \tag{368}$$

となり, 平均ゼロ, 分散 σ^2 の正規分布が得られた．

この方法を用いれば, 制約条件が増えて, 例えば「確率変数の高次のモーメントが一定である」という条件をも兼ね合わせた $P_X(x)$ を見つけなければならないときでも, その条件：

$$\int_{-\infty}^{\infty} dx\, x^n P_X(x) = m_n \tag{369}$$

を別の未定係数を用いて F の中に取り込み, 上述と同じ手続きをふめば良いということになる.

ここで示したように, この方法は割りと便利で重宝するので, 覚えておけば何かと役に立つことがあるかもしれない (⇒ 今回の **演習問題 11** 参照).

9.2 連続量の相互情報量

連続量のエントロピーについて見たので, 次に連続量の相互情報量について見ておこう. そこで, 相互情報量を次のように確率変数 X のエントロピー $H(X)$, 確率変数 Y のエントロピー $H(Y)$, 及び, X, Y の結合エントロピー $H(X,Y)$ を用いて

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{370}$$

と書き直しておく (なぜこの式が出るか, については各自が復習しておくように). このとき, x, y がそれぞれ離散値 $x_i, y_i\,(i = -\infty, \cdots, -1, 0, 1, \cdots, \infty)$ をとるものとし, $\Delta x \equiv x_{i+1} - x_i, \Delta y \equiv y_{i+1} - y_i$ と定義し, $\Delta x \to 0, \Delta y \to 0$ の極限を考えるものとする. このとき, (370) 式で定義される相互情報量は

$$\begin{aligned} I(X;Y) &= -\lim_{\Delta x \to 0} \sum_{i=-\infty}^{\infty} P_X(x_i)\Delta x \log\{P_X(x_i)\Delta x\} - \lim_{\Delta y \to 0} \sum_{i=-\infty}^{\infty} P_Y(y_i)\Delta y \log\{P_Y(y_i)\Delta y\} \\ &+ \lim_{\Delta x \to 0}\lim_{\Delta y \to 0} \sum_{i=-\infty}^{\infty}\sum_{i=-\infty}^{\infty} P_{XY}(x_i, y_i)\Delta x \Delta y \log\{P_{XY}(x_i, y_i)\Delta x \Delta y\} \end{aligned} \tag{371}$$

となるが, $P_{XY}(x_i, y_i)\Delta x \Delta y$ を y_i に関して和をとったものは $P_X(x_i)\Delta x$ に等しい, つまり

$$P_X(x_i)\Delta x = \sum_{i=-\infty}^{\infty} P_{XY}(x_i, y_i)\Delta x \Delta y \tag{372}$$

が成り立つこと等に注意すると (同時分布の周辺化), (371) 式は

$$\begin{aligned} I(X;Y) &= \lim_{\Delta x \to 0}\lim_{\Delta y \to 0} \sum_{i=-\infty}^{\infty}\sum_{i=-\infty}^{\infty} P_{XY}(x_i, y_i)\Delta x \Delta y \log\left\{\frac{P_{XY}(x_i, y_i)\Delta x \Delta y}{P_X(x_i)P_Y(y_i)\Delta x \Delta y}\right\} \\ &= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, P_{XY}(x,y) \log\frac{P_{XY}(x,y)}{P_X(x)P_Y(y)} \end{aligned} \tag{373}$$

となる. 従って, このように相互情報量を連続化する場合にはエントロピーのときに見られたような発散は生じない. これは相互情報量の形がエントロピーの差の形で書けるため, 発散を引き起こす部分が相殺されることによる (具体的には, 対数の中の分母・分子に共通な因子として $\Delta x \Delta y$ が現れ, これらが相殺することになる).

連続量の相互情報量について理解を深めるために, 次の例題を見ておこう.

例題 12

信号 x をある通信路に入力したところ, x には加法的雑音 ξ が加わって, 出力

$$y = x + \xi \tag{374}$$

が観測された. ここで ξ が平均ゼロ, 分散 σ_N^2 の正規分布：$P(\xi) = \mathrm{e}^{-\frac{\xi^2}{2\sigma_N^2}}/\sqrt{2\pi\sigma_N^2}$ に従うとすれば, (374) はこの通信路を特徴付ける条件付き確率 $P_{Y|X}(y|x)$ が

$$P_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi\sigma_N^2}}\,\mathrm{e}^{-\frac{(y-x)^2}{2\sigma_N^2}} \tag{375}$$

で与えられると言い換えることができることに注意しょう. さらに入力 x の従う確率分布 $P_X(x)$ に確率変数 X の 2 乗平均が σ^2 以下：

$$\int_{-\infty}^{\infty} dx\, x^2 P_X(x) \leq \sigma^2 \tag{376}$$

という条件を課すことにする. このとき以下の問いに答えよ.

(1) 条件付きエントロピー：

$$H(Y|X) = -\int_{-\infty}^{\infty} dx\, P_X(x) \int_{-\infty}^{\infty} dy\, P_{Y|X}(y|x) \log_2 P_{Y|X}(y|x) \tag{377}$$

を求めよ.

(2) 記号 $\langle \cdots \rangle$ を

$$\langle \cdots \rangle \equiv \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, (\cdots) P_{Y|X}(y|x) P_X(x) \tag{378}$$

で定義すると, 関係式：

$$\langle y^2 \rangle = \langle x^2 \rangle + \sigma_N^2 \leq \sigma^2 + \sigma_N^2 \tag{379}$$

が成り立つことを示せ.

(3) 上の (1)(2) と前に見た**例題 11** の結果を用いて

$$I(X;Y) = H(Y) - H(Y|X) \leq \frac{1}{2}\log_2\left(1 + \frac{\sigma^2}{\sigma_N^2}\right) \tag{380}$$

つまり, この通信路の通信路容量 C が

$$C = \frac{1}{2}\log_2\left(1 + \frac{\sigma^2}{\sigma_N^2}\right) \tag{381}$$

で与えられることを示せ.

(解答例)

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

$$\begin{aligned} H(Y|X=x) &= -\int_{-\infty}^{\infty} dy P_{Y|X}(y|x)\log P_{Y|X}(y|x) \\ &= -\int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi\sigma_N^2}}\,\mathrm{e}^{-\frac{(y-x)^2}{2\sigma_N^2}}\log\left\{\frac{1}{\sqrt{2\pi\sigma_N^2}}\,\mathrm{e}^{-\frac{(y-x)^2}{2\sigma_N^2}}\right\} \\ &= -\frac{\log\mathrm{e}}{2\sigma_N^2}\int_{-\infty}^{\infty}\frac{dy}{\sqrt{2\pi\sigma_N^2}}(y-x)^2\mathrm{e}^{-\frac{(y-x)^2}{2\sigma_N^2}} - \frac{1}{2}\log(2\pi\sigma_N^2) \\ &= -\frac{\log\mathrm{e}}{2\sigma_N^2}\int_{-\infty}^{\infty}\frac{dz}{\sqrt{2\pi\sigma_N^2}}\,z^2\mathrm{e}^{-\frac{z^2}{2\sigma_N^2}} - \frac{1}{2}\log(2\pi\sigma_N^2) \end{aligned} \tag{382}$$

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：

$$\int_{-\infty}^{\infty} dz\,z^2\,\mathrm{e}^{-az^2} = -\frac{\partial}{\partial a}\left\{\int_{-\infty}^{\infty} dz\,\mathrm{e}^{-az^2}\right\} \tag{383}$$

を評価することが必要となるが, **例題 11** の問題文中で示した公式 (345)：

$$\int_{-\infty}^{\infty} dz\,\mathrm{e}^{-az^2} = \sqrt{\frac{\pi}{a}} \tag{384}$$

を (383) 式の $\{\cdots\}$ に代入して, a に関する微分を実行すれば

$$\int_{-\infty}^{\infty} dz\,z^2\,\mathrm{e}^{-az^2} = \frac{\sqrt{\pi}}{2a\sqrt{a}} \tag{385}$$

が得られるから, (382) 式で $a=1/2\sigma_N^2$ とおくことにより, $H(Y|X=x)$ は

$$H(Y|X=x) = \frac{1}{2}\log\mathrm{e} + \frac{1}{2}\log(2\pi\sigma_N^2) = \frac{1}{2}\log(2\pi\mathrm{e}\sigma_N^2) \tag{386}$$

となる．この結果は 2 元対称通信路の場合と同様に, 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

$$\begin{aligned} H(Y|X) &= \int_{-\infty}^{\infty} dx P_X(x) H(Y|X=x) \\ &= \frac{1}{2}\log(2\pi\mathrm{e}\sigma_N^2)\int_{-\infty}^{\infty} dx P_X(x) = \frac{1}{2}\log(2\pi\mathrm{e}\sigma_N^2) \end{aligned} \tag{387}$$

となる．

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

$$\begin{aligned} \langle y^2\rangle &= \int_{-\infty}^{\infty} dx P_X(x)\int_{-\infty}^{\infty}\frac{y^2 dy}{\sqrt{2\pi\sigma_N^2}}\,\mathrm{e}^{-\frac{(y-x)^2}{2\sigma_N^2}} \\ &= \int_{-\infty}^{\infty} dx\,x^2 P_X(x)\int_{-\infty}^{\infty}\frac{dz}{\sqrt{2\pi\sigma_N^2}}\,\mathrm{e}^{-\frac{z^2}{2\sigma_N^2}} + \int_{-\infty}^{\infty}\frac{dz}{\sqrt{2\pi\sigma_N^2}}\,\mathrm{e}^{-\frac{z^2}{2\sigma_N^2}}\int_{-\infty}^{\infty} dx P_X(x) \\ &= \langle x^2\rangle + \sigma_N^2 \leq \sigma^2 + \sigma_N^2 \end{aligned} \tag{388}$$

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,

$$\int_{-\infty}^{\infty} dz\, z\, \mathrm{e}^{-az^2} = \left[-\frac{1}{2a}z^2\right]_{-\infty}^{\infty} = 0 \tag{389}$$

という事実を使い, 2 行目の右辺第 2 項の積分は (385) 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

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - \frac{1}{2}\log(2\pi\mathrm{e}\sigma_N^2) \tag{390}$$

であるから, あとは $H(Y)$ を最大化すればよいが, **例題 11** の結果 (356) 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

$$H(Y) \leq \frac{1}{2}\log(2\pi\mathrm{e}\langle y^2\rangle) \leq \frac{1}{2}\log(2\pi\mathrm{e}(\sigma^2+\sigma_N^2)) \tag{391}$$

と上から押さえられるので, 結局

$$I(X;Y) \leq \frac{1}{2}\log(2\pi\mathrm{e}(\sigma^2+\sigma_N^2)) - \frac{1}{2}\log(2\pi\mathrm{e}\sigma_N^2) = \frac{1}{2}\log\left(1+\frac{\sigma^2}{\sigma_N^2}\right) = C \tag{392}$$

が得られる．この C が正規雑音通信路の通信路容量である.

演習問題 11

ラグランジュの未定係数法を用いて, 確率変数 x が正の値のみをとり, x の平均値が m となるような確率分布 $P_X(x)$, つまり, 制約条件：

$$\int_0^{\infty} dx\, P_X(x) = 1, \quad \int_0^{\infty} dx\, x\, P_X(x) = m \tag{393}$$

を満たす確率分布 $P(x)$ の中で, 次で定義されるエントロピー：

$$H(X) = -\int_0^{\infty} dx\, P_X(x)\log_2 P_X(x) \tag{394}$$

を最大にするようなものを求めよ.

[**今後のスケジュール**]

7/11 (#12) 連続量の情報 II：標本化定理
7/25 (#13 最終回) 暗号を行けるとこまでやる.