情報理論 配布資料 #8

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 14 日

演習問題 7 の解答例

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：

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

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[1] が

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

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**：

$$H(X) = -\int_{-\infty}^{\infty} dx\, P_X(x)\log P_X(x) \tag{3}$$

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演習問題である. 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

$$P_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{x^2}{2\sigma^2}} \tag{4}$$

であり, この $P_X(x)$ に対し, エントロピー $H(X)$ は

$$\begin{aligned} H(X) &= -\int_{-\infty}^{\infty} dx\, P_X(x)\log\left\{\frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\right\} \\ &= -\log \mathrm{e}\int_{-\infty}^{\infty} dx\, P_X(x)\left\{-\frac{x^2}{2\sigma^2}\right\} + \frac{1}{2}\log(2\pi\sigma^2)\int_{-\infty}^{\infty} dx\, P_X(x) \\ &= \frac{1}{2}\log(2\pi\mathrm{e}\sigma^2) \end{aligned} \tag{5}$$

[1] 問題文ではこの「2 乗平均値」が「分散」と表記されていましたが, 必ずしも平均がゼロである必要はないので, ここでは 2 乗平均値が適切です.

となる．ここで, 条件 (2), 及び, 確率の規格化条件：

$$\int_{-\infty}^{\infty} dx\, P_X(x) \quad = \quad 1 \tag{6}$$

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

$$\frac{\log_2 \alpha}{\log_2 \mathrm{e}} = \log \alpha \tag{7}$$

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

$$\log \alpha \quad \leq \quad \alpha - 1 \tag{8}$$

が成り立つ．

(3) $H(x)$ と $\log_2(2\pi \mathrm{e}\sigma^2)/2$ の差を評価していこう．

$$\begin{aligned} H(X) - \frac{1}{2}\log_2(2\pi\mathrm{e}\sigma^2) &= -\int_{-\infty}^{\infty} dx\, P_X(x)\log P_X(x) + \int_{-\infty}^{\infty} dx\, P_X(x) \log\left\{\frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\right\} \\ &= \int_{-\infty}^{\infty} dx\, P_X(x) \log\left\{\frac{\mathrm{e}^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}P_X(x)}\right\} \end{aligned} \tag{9}$$

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

$$\begin{aligned} H(X) - \frac{1}{2}\log_2(2\pi\mathrm{e}\sigma^2) &\leq \int_{-\infty}^{\infty} dx\, P_x(X)\left\{\frac{\mathrm{e}^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}P_X(x)} - 1\right\} \\ &= \int_{-\infty}^{\infty} dx\, \frac{\mathrm{e}^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} - \int_{-\infty}^{\infty} dx\, P_X(x) = 0 \end{aligned} \tag{10}$$

従って

$$H(X) \quad \leq \quad \frac{1}{2}\log_2(2\pi\mathrm{e}\sigma^2) \tag{11}$$

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．

(参考)

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 $P_X(x)$ を求めたわけですが, 皆さんの中には「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」という人がいるかもしれません．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**というのをやりましたが, 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ます．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ますので, 参考としてそれを示しておきましょう．

まず, 条件 (6)(2) 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

$$\begin{aligned} F(P_X(x), \lambda, \beta) &= -\int_{-\infty}^{\infty} dx\, P_X(x) \log P_X(x) + \lambda \left\{ \int_{-\infty}^{\infty} dx\, P_X(x) - 1 \right\} \\ &+ \beta \left\{ \int_{-\infty}^{\infty} dx\, x^2 P_X(x) - \sigma^2 \right\} \end{aligned} \tag{12}$$

を作れば, 我々が最大化すべきなのは関数 F ということになります. 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 — になっている点です.

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

$$\frac{\delta F}{\delta P_X(x)} = -\int_{-\infty}^{\infty} dx\, \left\{ \log P_X(x) + 1 + \lambda + \beta x^2 \right\} = 0 \tag{13}$$

$$\frac{\partial F}{\partial \lambda} = \int_{-\infty}^{\infty} dx\, P_X(x) - 1 = 0 \tag{14}$$

$$\frac{\partial F}{\partial \beta} = \int_{-\infty}^{\infty} dx\, x^2 P_X(x) - \sigma^2 = 0 \tag{15}$$

が得られます. 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 を用いてあります.

さて, 上の 2,3 番目の式からは条件 (6)(2) が再度得られたことになります. 従って, まずは問題となるのが 1 番目の式ですが, 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

$$P_X(x) = \mathrm{e}^{-1-\lambda-\beta x^2} \tag{16}$$

が得られます. あとは λ, β を決めるのですが, この $P_X(x)$ を (14) に代入すれば

$$\mathrm{e}^{-1-\lambda} = \left(\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\beta x^2} \right)^{-1} \tag{17}$$

となります. これを用いて (16) を書き直すと

$$P_X(x) = \frac{\mathrm{e}^{-\beta x^2}}{\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\beta x^2}} \tag{18}$$

が得られますが, これを (15) に代入すると

$$\frac{\int_{-\infty}^{\infty} dx\, x^2 \mathrm{e}^{-\beta x^2}}{\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\beta x^2}} = \sigma^2 \tag{19}$$

すなわち

$$-\frac{\partial}{\partial \beta} \left\{ \log \int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\beta x^2} \right\} = \sigma^2 \tag{20}$$

が導かれます. 前回の演習問題に与えた積分公式から

$$\int_{-\infty}^{\infty} dx\, \mathrm{e}^{-\beta x^2} = \sqrt{\frac{\pi}{\beta}} \tag{21}$$

ですので, 結局 β が

$$\beta = \frac{1}{2\sigma^2} \tag{22}$$

と定まり, $P_X(x)$ は

$$P_X(x) = \frac{\mathrm{e}^{-\frac{x^2}{2\sigma^2}}}{\int_{-\infty}^{\infty} dx\,\mathrm{e}^{-\frac{x^2}{2\sigma^2}}} = \frac{1}{\sqrt{2\pi\sigma^2}}\,\mathrm{e}^{-\frac{x^2}{2\sigma^2}} \tag{23}$$

となり, 平均ゼロ, 分散 σ^2 の正規分布が得られました.

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 :

$$\int_{-\infty}^{\infty} dx\,x^n P_X(x) = m_n \tag{24}$$

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ります (簡単な練習問題なので時間のある人はやってみるように).

ここで示したように, この方法は割りと便利で重宝しますので, 覚えておけば何かと役に立つことがあるかもしれません.

演習問題 6 に関するコメント

こちらが思っていた以上に良くできていたのですが, 以下に気づいた点を 2,3 列記しておきます.

- $H(Y|X=1), H(Y|X=0)$ と $H(Y|X)$ は別物である. **前者はある特定の X の値に対する条件付きエントロピーであり, 後者は前者を X の確率分布 $P_X(x)$ で平均したもの (通常, 条件付きエントロピーと言った時にはこっちを指す)**. 問題 1. (1) で, これらを混同している人が数名いました.

- 通信路容量は

  $$C = \max_{P_X}\{H(Y) - H(Y|X)\}$$

  で定義され, 入力分布に関して, 相互情報量を最大化したものであり, **通信路の誤り確率 α に関して最大化したものではない**. この問題では誤り確率 α をある値に固定した上での容量を問うている. 当然, C は α の関数になっている. 問題 1.(2) で, C を求める際に α に関して最大化していた人が複数いました.

- 問題 2. (4) で **n 個の BSC を直列につなげた通信路の誤り確率は $\overline{\alpha}$ である**. α と $\overline{\alpha}$ の対応が見えていない人が多数いました. ちなみに, この問題が一番正答率が悪かった.

以上の点をもう一度確認しておいて下さい.

$\boxed{\textbf{演習問題 8}}$

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

$$y \quad = \quad x + \xi \tag{25}$$

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：

$$P(\xi) \quad = \quad \frac{1}{\sqrt{2\pi\sigma_N^2}} \mathrm{e}^{-\frac{\xi^2}{2\sigma_N^2}} \tag{26}$$

に従うとすれば, (25) は「この通信路を特徴付ける条件付き確率 $P_{Y|X}(y|x)$ が

$$P_{Y|X}(y|x) \quad = \quad \frac{1}{\sqrt{2\pi\sigma_N^2}} \mathrm{e}^{-\frac{(y-x)^2}{2\sigma_N^2}} \tag{27}$$

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：

$$\int_{-\infty}^{\infty} dx\, x^2 P_X(x) \quad \leq \quad \sigma^2 \tag{28}$$

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：

$$H(Y|X) \quad = \quad -\int_{-\infty}^{\infty} dx\, P_X(x) \int_{-\infty}^{\infty} dy\, P_{Y|X}(y|x) \log P_{Y|X}(y|x) \tag{29}$$

を求めよ.

(2) 記号 $\langle\cdots\rangle$ を

$$\langle\cdots\rangle \quad \equiv \quad \int_{-\infty}^{\infty} dxdy\, (\cdots) P_{Y|X}(y|x) P_X(x) \tag{30}$$

で定義すると, 関係式：

$$\langle y^2\rangle = \langle x^2\rangle + \sigma_N^2 \quad \leq \quad \sigma^2 + \sigma_N^2 \tag{31}$$

が成り立つことを示せ.

(3) 上の (1)(2) と前回の $\boxed{\textbf{演習問題 7}}$ の結果を用いて

$$I(X;Y) \quad = \quad H(Y) - H(Y|X) \leq \frac{1}{2}\log_2\left(1 + \frac{\sigma^2}{\sigma_N^2}\right) \tag{32}$$

つまり, この通信路の容量 C が

$$C \quad = \quad \frac{1}{2}\log_2\left(1 + \frac{\sigma^2}{\sigma_N^2}\right) \tag{33}$$

で与えられることを示せ.