数理解析研究所講究録
第1852巻 2013年 177-184

幾何分布のエントロピーとその周辺

湘南工科大学・工学部　落海　望
Nozomu Ochiumi
Faculty of Engineering, Shonan Institute of Technology

東京理科大学大学院・理学研究科・数理情報科学専攻　岸　祐太
Yuta Kishi
Department of Mathematical Science for Information Sciences
Graduate School of Science, Tokyo University of Science

1　はじめに

与えられたコスト

$$\boldsymbol{c}=(c_1,c_2,\ldots),\ 0\leqslant c_1\leqslant c_2\leqslant\ldots,\ (c_1<c_j \text{ for some } j)$$

に対して，制約条件（期待値固定）$\sum c_ip_i=\mu\ (\mu>c_1)$ の下でエントロピー $H(p_1,p_2,\ldots)=-\sum_{i\geqslant 1}p_i\log_2 p_i$ を最大にする分布が存在するとし $\boldsymbol{p}(\mu)=(p_1(\mu),p_2(\mu),\ldots)$ とする．これはボルツマン分布と呼ばれる．ラグランジュの未定乗数法より，ボルツマン分布は $p_i(\mu)=2^{-\lambda_0-\lambda c_i}$ の形をしていることが分かる．ここで，λ_0,λ は

$$\begin{cases}\displaystyle\sum_{i\geq 1}p_i(\mu)=\sum_{i\geq 1}2^{-\lambda_0-\lambda c_i}=1\\ \displaystyle\sum_{i\geq 1}c_ip_i(\mu)=\sum_{i\geq 1}c_i2^{-\lambda_0-\lambda c_i}=\mu\end{cases}$$

を満たし，$\mu=\sum_{i\geqslant 1}c_i\frac{2^{-\lambda c_i}}{\sum_{j\geqslant 1}2^{-\lambda c_j}}$ は λ に関して狭義単調減少関数である．よって，

$$Z(\lambda)=\sum 2^{-\lambda c_i}$$

とおくと，

$$p_i(\mu) = \frac{2^{-\lambda c_i}}{Z(\lambda)}$$

を得る．従って，ボルツマン分布のエントロピーは

$$H(\boldsymbol{p}(\mu)) = \lambda\mu + \log_2 Z(\lambda)$$

となる．$h(\mu) = H(\boldsymbol{p}(\mu))$ とおくと，$h(\mu)$ は μ に関して狭義単調増加，上凸関数である．よって，制約条件下での任意の分布に対するエントロピーに関して $h(\mu) \geqslant H(X)$ が成り立つことから $\mu \geqslant h^{-1}(H(X))$ を得る．しかし，一般のコストの場合，μ を $H(X)$ の陽の関数として表すこと，下から押さえることは困難であることが知られている．そこで本研究では具体的なコストについて個々に考えることにより，そのようなコストに対する期待値 μ の下界を $H(X)$ の関数で表すことを目的とする．以下，log の底は 2 とする．

2 具体的なコスト

2.1 重複があるコスト

2.1.1 重複度の導入

重複があるコスト；

$$\boldsymbol{c} = (c_1, c_2, \ldots) = (\underbrace{0, \ldots, 0}_{m_0}, \underbrace{1, \ldots, 1}_{m_1}, \underbrace{2, \ldots, 2}_{m_2}, \ldots),$$

$m_k \in \{0, 1, 2, \ldots\}$：コスト k の重複度

の場合を考える．このとき，ボルツマン分布のエントロピーは，

$$h(\mu) = -\sum_{k \geqslant 0} m_k \cdot (2^{-\lambda_0 - \lambda k}) \log(2^{-\lambda_0 - \lambda k}) = \lambda\mu + \lambda_0$$

と計算される．ここで，λ, λ_0 は

$$\begin{cases} \displaystyle\sum_{i \geq 1} p_i(\mu) = \sum_{k \geqslant 0} m_k 2^{-\lambda_0 - \lambda k} = 1 \\ \displaystyle\sum_{i \geq 1} c_i p_i(\mu) = \sum_{k \geqslant 0} k \cdot m_k 2^{-\lambda_0 - \lambda k} = \mu \end{cases}$$

を満たす．よって，

$$\sum_{k\geqslant 0} m_k k(2^{-\lambda})^k = \mu \sum_{k\geqslant 0} m_k (2^{-\lambda})^k$$

であるから，

$$g(x) = \sum_{k\geqslant 0} m_k x^k$$

とおくと，

$$xg'(x) = \mu g(x)$$

なる x の方程式の解が $2^{-\lambda}$ であり，$2^{\lambda_0} = g(2^{-\lambda})$ である．以下，具体的な重複度 m_k;

- $m_0 = 0,\ m_k = 1 \quad k = 1,2,3,\ldots \quad \boldsymbol{c} = (1,2,3,\ldots)$
- $m_k = 2^k \quad k = 0,1,2,\ldots \quad \boldsymbol{c} = (0,1,1,2,2,2,2,3,3,3,3,3,3,3,3,\ldots)$
- $m_0 = 0,\ m_k = 2^k \quad k = 1,2,3,\ldots \quad \boldsymbol{c} = (1,1,2,2,2,2,3,3,3,3,3,3,3,3,\ldots)$
- $m_k = k \quad k = 0,1,2,\ldots \quad \boldsymbol{c} = (1,2,2,3,3,3,\ldots)$

を与えることにより $g(x)$ を陽な関数で表せる場合を考える．

2.1.2 $m_0 = 0,\ m_k = 1,\ k = 1,2,3,\ldots$ の場合

重複度が $m_0 = 0,\ m_k = 1,\ k = 1,2,3,\ldots$ の場合を考える．このとき，コスト $\boldsymbol{c} = (1,2,3,\ldots)$ である．この場合は大変興味深く，次のような推測問題として知られている [1, 2]．ある確率に従い $\mathcal{X} = \{x_1, x_2, \ldots\}$ に値をとる確率変数 X に対して，"Yes" が出るまで，"x_i ですか？" と質問していく．この問題は例えば，何らかの暗号解読法によって可能性が絞られたのち，暗号解読者が 1 回ずつ秘密鍵を試さなければならない状況などに表れる．最適な，すなわち平均質問回数が最小となる推測の戦略は明らかに，確率の高いものから順に推測してくことである．以下，$\boldsymbol{p} = (p_1, p_2, \ldots)$ は $p_1 \geqslant p_2 \geqslant \ldots$ を満たすとする．このとき，平均質問回数は $\mu = \sum_{i\geqslant 1} i p_i$ によって与えられる．制約条件（平均質問回数固定）$\mu = \sum_{i\geqslant 1} i p_i$ の下でのボルツマン分布 $\boldsymbol{p}(\mu)$ は

$$g(x) = \frac{x}{1-x}\ ,\ 2^{\lambda} = \frac{\mu}{\mu - 1}\ ,\ 2^{\lambda_0} = \mu - 1$$

を満たす．このとき $\boldsymbol{p}(\mu)$ はパラメータ $\frac{1}{\mu}$ の幾何分布 $p_i(\mu) = \frac{1}{\mu}\left(1 - \frac{1}{\mu}\right)^{i-1}$ となる．幾何分布のエントロピー（$h_G(\mu)$ とする）は

$$h_G(\mu) = \log(\mu - 1) + \log\left(1 - \frac{1}{\mu}\right)^{-\mu}$$

となる．Massey [1] は右辺第 2 項を定数で評価することにより μ の下界を得た．

定理 A ([1]). $H(X) \geqslant 2$ に対して,

$$\mu \geqslant \frac{1}{4}2^{H(X)} + 1.$$

さらに，落海-柳田 [2] では,

$$2^{h_G(\mu)} = (\mu - 1)\left(1 - \frac{1}{\mu}\right)^{-\mu}$$

が μ に関して上凸関数であることを示し，任意の $a > 1, \mu > 1$ に対して,

$$2^{H(X)} \leqslant 2^{h_G(\mu)} = f(\mu) \leqslant f'(a)(\mu - a) + f(a)$$

であることから次の定理を得た．ここで $f(x) = (x-1)\left(1 - \frac{1}{x}\right)^{-x}$ とする.

定理 B ([2]). 任意の $a > 1$ に対して,

$$\mu \geqslant \frac{1}{f'(a)}2^{H(X)} + \left(a - \frac{f(a)}{f'(a)}\right)$$

が成り立つ.

さらに本研究では，コスト c_1 の確率が既知の場合（$p_1 = p$ とする）について考える．エントロピーの分枝性より

$$H(X) = H(p, p_2, p_3, \dots) = H(p, 1-p) + (1-p)H\left(\frac{p_2}{1-p}, \frac{p_3}{1-p}, \cdots\right)$$

となるので，制約条件（平均質問回数固定）$\mu = \sum_{i \geqslant 2} i p_i$ の下で $H\left(\frac{p_2}{1-p}, \frac{p_3}{1-p}, \cdots\right)$ を最大にする分布 $\left(\frac{p_2(\mu)}{1-p}, \frac{p_3(\mu)}{1-p}, \dots\right)$ を求めればよい．すると，パラメータ $\frac{1-p}{\mu-1}$ の幾何分布となる．従って,

$$H\left(\frac{p_2}{1-p}, \frac{p_3}{1-p}, \cdots\right) \leqslant h_G\left(\frac{\mu - 1}{1-p}\right)$$

となる．以上から，定理 B と同様にして以下の定理を得る.

定理 1 (O-K). 任意の $a > 1$ に対して,

$$\mu \geqslant \frac{p^p(1-p)^{1-p}}{g'(a)}2^{H(X)} + \left(a - \frac{g(a)}{g'(a)}\right)$$

が成り立つ．$g(x) = \left(\frac{1}{1-p}\right)^{1-p}(x-1)^{x-1}(x+p-2)^{2-x-p}$ とする.

2.1.3 $m_k = 2^k,\ k = 0, 1, 2, \ldots$ の場合

重複度が $m_k = 2^k,\ k = 0, 1, 2, \ldots$ の場合を考える．このときコスト $\boldsymbol{c} = (0, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, \ldots)$ である．この場合も大変興味深い問題である [3, 4, 5]．このコストは情報源符号化における 2 元 1 対 1 符号（空語あり）の符号語の符号長とみなすことができる．このとき，制約条件（平均符号長固定）$\mu = \sum_{i \geqslant 1} c_i p_i$ の下でのボルツマン分布 $\boldsymbol{p}(\mu)$ は

$$g(x) = \frac{1}{1-2x},\ 2^{\lambda} = \frac{2(\mu+1)}{\mu},\ 2^{\lambda_0} = \mu + 1$$

を満たす．このボルツマン分布のエントロピーは幾何分布のエントロピーで次のように表現される．

$$h(\mu) = \mu + h_G(\mu + 1)$$

従って，幾何分布のエントロピーの上凸性より

$$H(X) \leqslant h(\mu) = \mu + h_G(\mu+1) \leqslant \mu + h_G{}'(a+1)(\mu - a) + h_G(a+1)$$

となり，以下の定理を得る．

定理 2 ([5], O-K). 任意の $a > 0$ に対して，

$$\mu \geqslant \frac{H(X) - \log(a+1)}{\log \frac{a+1}{a} + 1}$$

が成り立つ．

また，重複度が $m_0 = 0,\ m_k = 2^k,\ k = 1, 2, 3, \ldots$ の場合のコスト $\boldsymbol{c} = (1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, \ldots)$ は，空語なしの 1 対 1 符号の符号語の符号長となっている．そのとき，平均符号長 μ の下界は定理 2 と同様にして以下で与えられる．

定理 3 ([5], O-K). 任意の $a > 1$ に対して，

$$\mu \geqslant \frac{H(X) - \log(a-1)}{\log \frac{a}{a-1} + 1}$$

が成り立つ．

2.1.4 $m_k = k,\ k = 0, 1, 2, \ldots$ の場合

重複度が $m_k = k,\ k = 0, 1, 2, \ldots$ の場合を考える．このとき，コスト $\boldsymbol{c} = (1, 2, 2, 3, 3, 3, 4, 4, 4, 4, \ldots)$ である．制約条件（期待値固定）$\mu = \sum_{i \geqslant 1} c_i p_i$ の下でのボルツマン分布 $\boldsymbol{p}(\mu)$ は

$$g(x) = \frac{x}{(1-x)^2},\ 2^{\lambda} = \frac{\mu+1}{\mu-1},\ 2^{\lambda_0} = \frac{(\mu+1)(\mu-1)}{4}$$

を満たす．このボルツマン分布のエントロピーは幾何分布のエントロピーで次のように表現される．

$$h(\mu) = 2h_G\left(\frac{\mu+1}{2}\right)$$

従って，幾何分布のエントロピーの上凸性より，

$$H(X) \leqslant h(\mu) = 2h_G\left(\frac{\mu+1}{2}\right) \leqslant 2\left(h_G{}'\left(\frac{a+1}{2}\right)(\mu-a) + h_G\left(\frac{a+1}{2}\right)\right)$$

となり，以下の定理を得る．

定理 4 (O-K)．任意の $a > 1$ に対して，

$$\mu \geqslant \frac{H(X) + \log \frac{4}{(a+1)(a-1)}}{\log \frac{a+1}{a-1}}$$

が成り立つ．

また，$2^{h(\mu)} = 2^{2h_G(\frac{\mu+1}{2})}$ は，任意の $a > 1, \mu > 1$ に対して，

$$2^{H(X)} \leqslant 2^{h(\mu)} = \left(2^{h_G(\frac{\mu+1}{2})}\right)^2 \leqslant (l'(a)(\mu-a) + l(a))^2$$

となる．ここで $l(x) = \left(\frac{x+1}{2} - 1\right)\left(1 - \frac{2}{x+1}\right)^{-\frac{x+1}{2}}$ とする．よって以下の定理を得る．

定理 5 (O-K)．任意の $a > 1$ に対して，

$$\mu \geqslant \frac{1}{l'(a)} 2^{\frac{H(X)}{2}} + \left(a - \frac{l(a)}{l'(a)}\right)$$

が成り立つ．

2.2 $c_i = i^2$ の場合

$c_i = i^2$ の場合を考える．このときコスト $\boldsymbol{c} = (1, 4, 9, \ldots)$ となる．この場合，このコストの期待値は 2.1.2 で述べた推測問題における 2 次積率となっている．制約条件（期待値固定）$\mu = \sum_{i \geqslant 1} i^2 p_i$ の下でのボルツマン分布 $\boldsymbol{p}(\mu)$ は

$$Z(\lambda) = \sum 2^{-\lambda i^2}$$

とおくと，

$$p_i(\mu) = \frac{2^{-\lambda i^2}}{Z(\lambda)}$$

となる．しかし，ボルツマン分布のエントロピー

$$H(\boldsymbol{p}(\mu)) = \lambda\mu + \log Z(\lambda)$$

を μ の式で表すためには

$$\frac{\sum_{i \geqslant 1} i^2 2^{-\lambda i^2}}{\sum_{i \geqslant 1} 2^{-\lambda i^2}} = \mu$$

を λ について解かなければならない．そのうえで，ボルツマン分布のエントロピー $h(\mu)$ に関して，

$$2^{h(\mu)} = Z(\lambda) 2^{\lambda\mu}$$

が上凸関数であること，つまり，

$$\frac{d^2}{d\mu^2} 2^{h(\mu)} = 2^{h(\mu)} \left((\lambda \ln 2)^2 - \frac{1}{\sigma^2} \right)$$

が負であることが示されれば，2.1.2 で述べた推測問題における分散に関しての結果が得られることが期待される．ただし，$\sigma^2 = \sum (i^2)^2 p_i(\mu) - (\sum i^2 p_i(\mu))^2$ である．

参考文献

[1] J. L. Massey, "Guessing and Entropy," *Proc. IEEE Int. Symp. on Info. Th.* (1994), 204.

[2] 落海望, 柳田昌宏, "An Improved Bound in Guessing," 第 29 回情報理論とその応用シンポジウム, 情報理論とその応用学会, (2006), 67.

[3] N. Alon, A. Orlitsky, "A lower bound on the expected length of one-to-one codes." *IEEE Trans. Inform. Theory* **40**, (1994), 1670–1672.

[4] C. Blundo, R. De Prisco, "New bounds on the expected length of one-to-one codes." *IEEE Trans. Inform. Theory* **42**, (1996), 246–250.

[5] J. Cheng, T. Huang, C. Weidmann, "New bounds on the expected length of optimal one-to-one codes." *IEEE Trans. Inform. Theory* **53**, (2007), 1884–1895.