２０１９年5月6日（川端）

データ圧縮基礎資料

6 AEP 性 (漸近等分割性)

補題 1 (チェビシェフの不等式) 期待値 $\mu = EX$, 分散 $\sigma^2 = V(X) = E(X-\mu)^2$ をもつ確率変数 X について，チェビシェフの不等式：

$$P\Big(|X-\mu| \geq k\sigma\Big) \leq \frac{1}{k^2}$$

が成り立つ．

証明：

$$\begin{aligned}
\text{左辺} &= E1\{|X-\mu| \geq k\sigma\} \\
&\leq E\frac{(X-\mu)^2}{k^2\sigma^2} \\
&= \frac{1}{k^2}.
\end{aligned} \tag{1}$$

定理 1 無記憶情報源 $\{Y_l\}$ において，Y_1 の期待値，分散が有限のとき，その平均値 $\frac{1}{n}\sum_{l=1}^{n} Y_l$ は期待値に確率収束する．言い換えると，あらゆる $\varepsilon > 0$ において，

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{l=1}^{n} Y_l - EY_1\right| \geq \varepsilon\right) = 0$$

が成り立つ．

証明：チェビシェフの不等式を用い，$\varepsilon = k\sqrt{\frac{1}{n}V(Y)}$ とおくと，

$$P\left(\left|\frac{1}{n}\sum_{l=1}^{n} Y_l - EY_1\right| \geq \varepsilon\right) \quad = \frac{V(Y)}{n\varepsilon^2} \tag{2}$$

が成り立つ．この右辺は $n \to \infty$ のとき零に収束する．

定理 2 $\mathcal{X}$ 上の離散無記憶情報源 $\{X_l\}$ において，$H(X) < \infty$, $V(-\log p(X)) < \infty$ のとき，$-(1/n)\log p(X^n)$ は $H(X)$ に確率収束する．

注意 1 同じ条件下で，もっと強い収束：

$$P\left(\lim_{n\to\infty} -\frac{1}{n}\log p(X^n) = H(p())\right) = 1 \tag{3}$$

がいえる．確率変数列のこの収束を概収束といい，

$$\lim_{n\to\infty} -\frac{1}{n}\log p(X^n) = H(p()) \ \ (a.s.)$$

のようにかく．

定義 1 確率分布 $p : \mathcal{X} \to \mathcal{R}$ により定義される，無記憶情報源について，集合：

$$\{x^n : \left| -\frac{1}{n}\log p(x^n) - H(p())\right| \leq \varepsilon\}$$

をこの確率分布に関する ε-標準集合とよび，$A_\varepsilon^{(n)}$ とかくまたこれに属する系列 x^n を ε 標準系列という．

次の定理がなり立つ．

定理 3 (AEP 性) $p : \mathcal{X} \to \mathcal{R}$ に関する，$A_\varepsilon^{(n)}$ について以下が成り立つ．

1.

$$\begin{aligned} x^n \in A_\varepsilon^{(n)} \leftrightarrow 2^{-n(H(X^n)+\varepsilon)} &\leq p(x^n) \\ &\leq 2^{-n(H(X^n)-\varepsilon)} \end{aligned}$$

2. $\exists n_0(\varepsilon), \forall n \geq n_0(\varepsilon)$,

$$P\Big(X^n \in A_\varepsilon^{(n)}\Big) > 1 - \varepsilon$$

3.

$$\sharp A_\varepsilon^{(n)} \leq 2^{n(H(X)+\varepsilon)}$$

4. $\exists n_0(\varepsilon), \forall n \geq n_0(\varepsilon)$,

$$\sharp A_\varepsilon^{(n)} \geq (1-\varepsilon) 2^{n(H(X)-\varepsilon)}.$$

注意 2 *AEP* 性は *Asymptotic Equipartition Property* の略である．$\mathcal{X}^n$ はほぼ確率 1 の標準集合と確率的に無視可能な部分に分かれるが，標準集合において，系列は「一様分布」している（あるいは「漸近的に等分割」している）．

証明：1 は標準系列の定義のいい替えである．2 は，定理**??**そのものである．3．は不等式

$$\begin{aligned} 1 &= \sum_{x^n \in \mathcal{X}^n} p(x^n) \\ &\geq \sum_{x^n \in A_\varepsilon^{(n)}} p(x^n) \\ &\geq \sum_{x^n \in A_\varepsilon^{(n)}} 2^{-n(H(X^n)+\varepsilon)} \quad (1.\text{より}) \\ &= \sharp A_\varepsilon^{(n)} 2^{-n(H(X^n)+\varepsilon)} \end{aligned}$$

からいえる．4．については，2 を成り立たせる $n \geq n_0(\varepsilon)$ について，

$$\begin{aligned} 1 &= \sum_{x^n \in \mathcal{X}^n} p(x^n) \\ &= \sum_{x^n \in A_\varepsilon^{(n)}} p(x^n) + \sum_{x^n \notin A_\varepsilon^{(n)}} p(x^n) \\ &\leq 2^{-n(H(X^n)-\varepsilon)} \Bigg(\sum_{x^n \in A_\varepsilon^{(n)}} 1 \Bigg) + P\Big(X^n \notin A_\varepsilon^{(n)}\Big) \\ &\qquad (1.\text{より}) \\ &\leq 2^{-n(H(X^n)-\varepsilon)} \sharp A_\varepsilon^{(n)} + \varepsilon \qquad (2.\text{より}) \end{aligned}$$

がなりたつが，この不等式を変形することにより得られる．さて，情報理論では多出力情報源や，ネットワーク型通信路の符号化の問題を考えることができるが，AEP 性は複雑な問題の本質をえぐりだすのに有用である．例えば，AEP 性を用いると情報源符号化の順定理が簡単に得られる．

定理 4 有限のアルファベットをもつ定常無記憶情報源 $\{X_l\}$ を考える．これに対して，情報源ブロックサイズ n の一意復号可能な可変長符号の列が存在し，$n \to \infty$ においてその期待平均符号長は情報源のエントロピーに収束する．

証明：$x^n \in \mathcal{X}^n$ にたいして一意復号可能な符号 φ_n を以下のように定義する：$x^n \in A_\varepsilon^{(n)}$ のとき，フラグビット 1 に x^n の $A_\varepsilon^{(n)}$ における固定符号長 $\lceil \log \sharp A_\varepsilon^{(n)} \rceil$ の数え上げ符号を連接する．$x^n \notin A_\varepsilon^{(n)}$ のとき，フラグビット 0 に x^n の $\mathcal{X}^n$ における固定符号長 $\lceil \log \sharp \mathcal{X}^n \rceil$ の数え上げ符号を連接する．任意の $\varepsilon > 0$ にたいして，$n_0(\varepsilon/2)$ を定理のようにとる．$n \geq n_0(\varepsilon/2)$ にたいして，期待平均符号長は，

$$
\begin{aligned}
E\varphi(x^n) \text{ の長さ} &\leq P\Big(X^n \in A_\varepsilon^{(n)}\Big)(2 + \log \sharp A_\varepsilon^{(n)}) + \\
&\quad P\Big(X^n \notin A_\varepsilon^{(n)}\Big)(2 + \log \sharp \mathcal{X}^n) \\
&\leq 2 + \log \sharp A_\varepsilon^{(n)}) + \varepsilon/2(2 + \log \sharp \mathcal{X}^n) \\
&\leq 2 + n(H(X) + \varepsilon/2) + \varepsilon/2(2 + n \log \sharp \mathcal{X})
\end{aligned}
$$

のように上界できるが，両辺を n で割り，適当な $n_1(\varepsilon) \geq n_0(\varepsilon/2)$ をとるならば，$n \geq n_1(\varepsilon)$ にたいして，右辺は $H(X) + \varepsilon$ で上界できる．一方一意復号可能な符号であるから，平均期待符号長は $H(X)$ で下界される．