ベイズで、2値変量の間の条件付き確率を推定する

鈴木 譲

大阪大学大学院基礎工学研究科

まず、長さ n の $X=0,1$ の各2進系列 x^n に対して、a を正の定数、c を1の頻度、$w(\theta)=K\theta^{a-1}(1-\theta)^{a-1}$ (K: 定数) として、$Q(x^n):=\int_0^1 \theta^c(1-\theta)^{n-c}w(\theta)d\theta$ の値を割り当てる。この値は、$Q(x^n)\geq 0$ および $\sum_{x^n}Q(x^n)=1$ を満足していて、系列の確率のような解釈ができる。次に、別の長さ n の各2進系列 y^n について、$Y=0,1$ のもとでの $X=0,1$ の条件付確率に相当する何らかの $Q(x^n|y^n)$ を計算して、$Q(x^n)$ と $Q(x^n|y^n)$ の大小を比較して、独立性を検定することを検討する。

この場合、$Q(x^n):=\prod_{i=1}^n \frac{c_{i-1}(x_i)+a}{i-1+2a}$ および $Q(x^n|y^n):=\prod_{i:y_i=0}\frac{c_{i-1}(x_i,y_i)+b}{c_{i-1}(y_i)+2b}\cdot\prod_{i:y_i=1}\frac{c_{i-1}(x_i,y_i)+b}{c_{i-1}(y_i)+2b}$ で計算する方法と、$Q(x^n,y^n):=\prod_{i=1}^n\frac{c_{i-1}(x_i,y_i)+a}{c_{i-1}(y_i)+4a}$ および $Q(x^n|y^n):=\frac{Q(x^n,y^n)}{Q(y^n)}$ で計算する方法がある。ただし、$c_{i-1}(x_i)$ で $(x_1,\cdots,x_{i-1})$ における x_i の頻度、$c_{i-1}(x_i,y_i)$ で $(x_1,\cdots,x_{i-1}),(y_1,\cdots,y_{i-1})$ における (x_i,y_i) の頻度とした。以下では、上記の $Q(x^n)$ および $Q(x^n,y^n)$ をそれぞれ、$f(x^n,a)$ および $g(x^n,y^n,a)$ とかくものとする。ここで、前者の場合、$Q(x^n|y^n)Q(y^n)=Q(y^n|x^n)Q(x^n)$ とするためには、$a=2b$ が必要になり、$Q(x^n|y^n)Q(y^n)=g(x^n,y^n,a/2)$ となる (Bayesian Dirichlet equivalent uniform [2], BDeu)。また、後者 (Quotient Jeffreys', QJ) の場合、$a=1/2$ (Jeffreys の事前確率) をおくことが多い。したがって、独立性検定は、独立である事前確率を 1/2 とすれば、それぞれ、$f(x^n,a)f(y^n,a)\geq g(x^n,y^n,a/2)$ および $f(x^n,1/2)f(y^n,1/2)\geq g(x^n,y^n,1/2)$ であるときに、独立であるとみなすことになる。2値でなく多値でも、2変数でなく多変数でも、独立性だけでなくベイジアンネットワークの構造学習などの条件付き独立性の検定にも適用される [1])。

本論文では、BDeu と QJ の性能の比較を行い、後者の卓越性に関して2種類の命題を証明する。まず、$x_i=h(y_i)$, $i=1,\cdots,n$ なる関数 $h:\{0,1\}\to\{0,1\}$ が存在するとき、確率変数 Z に関する系列 z^n が利用可能であっても、X の Y に関する経験的条件付エントロピーが0であれば、パラメータ数のより少ないモデルが選択されるべきであろう [3]。本論文では、QJ では、X が Y のみに依存するモデルを選択するが、BDeu では、X が Y,Z の両方に依存するモデルを選択する (**定理 1**) ことを証明した。

また、真のパラメータ (確率、条件付確率の値) を固定して、n を大きくすると、両者とも誤り率が0になる、すなわち一致性を満足する。しかし、BDeu では n をいくら大きくしても、誤り率が一定以上のパラメータの組が存在し (pointwise な一致性)、QJ では各 n で最も誤り率の高いパラメータの誤り率が、0に収束する (uniform な一致性)。そのような証明を行った (**定理 2**)。

参考文献

[1] J. Suzuki. “A novel Chow-Liu algorithm and its application to gene differential analysis”. *International Journal of Approximate Reasoning*, 2017.

[2] W. Buntine. Theory refinement on Bayesian networks. In *Uncertainty in Artificial Intelligence*, pages 52–60, Los Angels, CA, 1991.

[3] J. Suzuki and J. Kawahara. “Branch and Bound for Regular Bayesian network structure learning”. *Uncertainty in Artificial Intelligence*, Sydney, Australia, August, 2017.