

いま, \mathbf{X}, \mathbf{S} が直線線分 V 上に配置するとして, $|\mathcal{Q}|=2\pi$, $f_\theta(\mathbf{s}_i, \mathbf{s}_{i+1}) \equiv f_\theta(\mathbf{s}_i \cdot \mathbf{s}_{i+1})$ の場合を考える
と, $\mathbf{s}_i \cdot \mathbf{s}_{i+1} = \cos(\phi_{i+1} - \phi_i)$ に注意して,

$$\begin{aligned} Z(f_\theta; N) &= \frac{1}{(2\pi)^N} \int_0^{2\pi} d\phi_1 \prod_{i=1}^{N-1} \int_0^{2\pi} \exp\{-f_\theta(\cos\phi_{i+1})\} d\phi_{i+1} \\ &= \left[\frac{1}{2\pi} \int_0^{2\pi} \exp\{-f_\theta(\cos\phi)\} d\phi \right]^{N-1} \end{aligned}$$

となり, 規格化因子 Z が一重積分のみで表される(ここで \mathbf{s}_{i+1} の角 ϕ_{i+1} を \mathbf{s}_i を始線にして定めた).
これから数値積分によって対数尤度は容易に計算できる.

報告では, $f_\theta(\mathbf{s}_i \cdot \mathbf{s}_{i+1}) \equiv -\theta \mathbf{s}_i \cdot \mathbf{s}_{i+1}$ の場合(一次元 XY モデル)に対する計算機シミュレーション
を行ない, 母数 θ の最尤推定を行なって真の母数値と良く一致すること, およびその標本誤差の実験値
が理論値と良く一致することを示した. そして実データの解析例として, ヒトデの移動データを解析し
て系列相関分析との定性的な一致を見た. また講演では, 二次元 Ising モデルに対する厳密解のこの問題
への適用可能性についても述べた.

連結ベクトルの分布とその応用

馬場 康 維

1. 連結ベクトル表現

k 個の状態 L_1, L_2, \dots, L_k で表現される系を考える. 状態 L_j をベクトルの方向 θ_j で表現することに
する. 状態 L_j に属する標本をベクトル

$$\mathbf{x}_j = (w_j \cos \theta_j, w_j \sin \theta_j) \quad (0 \leq \theta_j \leq \pi)$$

で表現し, このベクトルを連結すればデータの構造を表現できる. これを連結ベクトル表現という. こ
こでは連結ベクトルの分布(Baba (1989))とその分布を用いる方法について述べる.

2. 合成ベクトルの座標軸

k 個の状態に対応するベクトルの合成ベクトルを

$$\mathbf{y} = \sum_{j=1}^k \mathbf{x}_j = (\sum w_j \cos \theta_j, \sum w_j \sin \theta_j)$$

とする. \mathbf{y} は 2 次元のベクトルであるから, 2 次元の量を表現できる. そこで \mathbf{y} の座標系について考え
る. ここで w_j に対して

$$\sum w_j = 1$$

という条件をつけておく.

$$c = \sum w_j \cos \theta_j, \quad s = \sum w_j \sin \theta_j$$

とすると

$$\mathbf{y} = \sum \mathbf{x}_j = (c, s)$$

である. \mathbf{y} の大きさと方向をそれぞれ R, ϕ とし,

$$\mathbf{y} = (R \cos \phi, R \sin \phi)$$

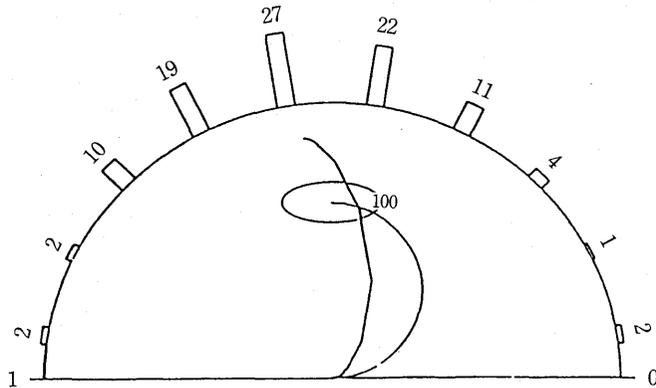


図1. 100個の標本の連結ベクトル表現, 一様分布の母集団曲線, 棄却楕円.

と表わすことにすると

$$\phi = \arctan(s/c) \quad (0 \leq \phi \leq \pi)$$

$$R = \sqrt{c^2 + s^2} = \sum w_j \cos(\theta_j - \phi)$$

となる. ϕ は角度 θ_j の重みつき平均方向であり,

$$V_c = 1 - R$$

は重みつき円分散である. $\phi, 2V_c$ は θ_j の範囲が狭いときは通常の意味と分散に近づくことが知られている. したがって合成ベクトルの座標系は平均と分散を表わす座標系である.

3. 合成ベクトルの漸近分布

n 個の標本をとったとして各状態に属する標本数を, f_1, f_2, \dots, f_k とする. $p_j = E(f_j)$ すなわち p_j を L_j の占有確率とする. f_1, f_2, \dots, f_k を多項分布

$$\frac{n!}{f_1! f_2! \dots f_k!}$$

にしたがう確率変数であると見做すと, (f_1, f_2, \dots, f_k) の漸近分布は正規分布になる.

したがって (c, s) の漸近分布は以下の平均ベクトルと分散共分散行列を持つ正規分布になる.

$$\mu = [\sum p_j \cos \theta_j, \sum p_j \sin \theta_j]$$

$$\Sigma = \frac{1}{n^2} \begin{bmatrix} \sum \sum C_\alpha C_\beta \sigma_{\alpha\beta} & \sum \sum C_\alpha S_\beta \sigma_{\alpha\beta} \\ \sum \sum S_\alpha C_\beta \sigma_{\alpha\beta} & \sum \sum S_\alpha S_\beta \sigma_{\alpha\beta} \end{bmatrix}$$

ここで $\sigma_{\alpha\beta} = p_\alpha \delta_{\alpha\beta} - n p_\alpha p_\beta$, $C_\alpha = \cos \theta_\alpha$, $S_\alpha = \sin \theta_\alpha$ である.

4. 漸近分布の応用

"順位グラフ" によるグラフ検定はこの漸近分布を用いた例である(Baba (1986)). ここでは標本分布のグラフ表現と分布に関するグラフ検定について述べる.

連続型分布の n 個の標本が k 個の階級に分けられているとする. この場合 L_1, L_2, \dots, L_k は階級に対応する. 標本が階級 j とされた頻度を f_j とする. θ を階級値に対応させ,

$$w_j = f_j/n, \quad \theta_j = \pi \left(j - \frac{1}{2} \right) / k$$

とすればヒストグラムのかわりに分布を表わす折れ線が描ける。また、漸近分布から得られる棄却楕円を描くことにより平均および分散に関する仮説検定ができる。

たとえば母集団が一様分布

$$f(\theta) = \frac{1}{\pi} \quad (0 < \theta < \pi)$$

という仮説を検定する場合を考えよう。このときは

$$\begin{aligned} \mu &= (0, \mu_0) \\ \Sigma &= \begin{bmatrix} \frac{1}{2n} & 0 \\ 0 & \frac{1}{2n}(1 - 2\mu_0) \end{bmatrix} \\ \mu_0 &= \frac{1}{k} \operatorname{cosec}(\pi/2k) \end{aligned}$$

である。これをもとに、図1に $k=10$, $n=100$ のときの棄却楕円を描いておいた。これは仮説が真のとき合成ベクトルの終点の5%がこの領域の外に出ることを示す楕円である。半円の上の柱は $n=100$ の標本の分布を示しており、いわばヒストグラムに対応する。この標本に対応する折れ線の終点は楕円の外に出ているから、仮説は棄却されることになる。

様々な母集団分布に対し漸近分布を求め、連結ベクトルとの距離を求めれば、分布の選択ができる。しかしながら、座標系が平均と分散であることから、この方法は平均と分散が同じ分布には無効である。ところで、 $n \rightarrow \infty$, $k \rightarrow \infty$ とし

$$f_j/n \rightarrow f(\theta)$$

という場合を考えてみよう。

$$\begin{aligned} c(t) &= \int_{-\infty}^t f(\theta) \cos \theta \, d\theta \\ s(t) &= \int_{-\infty}^t f(\theta) \sin \theta \, d\theta \\ \mathbf{y}(t) &= (c(t), s(t)) \end{aligned}$$

とおき $\mathbf{y}(t)$ の軌跡を描けば母集団分布を表わす滑らかな曲線が得られる。したがって、母集団を表わす曲線と連結ベクトルを比較することによって分布の選択ができる。

参 考 文 献

- Baba, Y. (1986). Graphical analysis of rank data, *Behaviormetrika*, **19**, 1-15.
 Baba, Y. (1989). Data analysis based on direction, Symposium on the Analysis of Statistical Information, Tokyo, 259-274.