

継続調査と平滑化法による世論調査 結果からの情報抽出

— ある事例研究 —

統計数理研究所 柏 木 宣 久

(1993 年 11 月 受付)

1. はじめに

我国では、政府、報道機関、調査機関等々により、多くの世論調査が実施されている。それら世論調査で取り扱われる話題は多岐にわたっており、調査という手法は、話題に依らず、世論を知るための一般的な方法として定着している。ただし、個々の話題についてみると、一過性の問題として扱われる場合が多く、継続的に調査されている話題は一部に過ぎない。話題の適時性や費用などの問題もあり、継続調査が必ずしも常に現実的とは思われないが、最近の調査環境の悪化に伴う回収率の低下や人間の反応の多様性および複雑さを考えた場合、継続調査は世論調査結果から情報を抽出するために重要である。本稿では、平滑化法による世論調査結果の解析を通じ、継続調査の有用性を示す。

解析で用いる平滑化法は、階差制約とベイズ型尤度に基づくベイズ的平滑化法である。その基本的枠組み、特にベイズ型尤度の利用は、Akaike (1980) により議論されている。ベイズ的平滑化法を用いるのは、他の方法、すなわち、核関数による回帰や平滑化スプラインなどの分布を仮定しない方法、および解析関数による回帰や回帰スプラインなどの尤度に基づく方法に比べ、滑らかさの調整の容易さや推論の組立て易さの点において優れているからである（詳細については、柏木・柳本 (1990) を参照されたい）。現在までに、ベイズ型尤度に基づく多くの方法が開発されている（例えば、柏木 (1982)、柏木・伊谷 (1986)、Kashiwagi and Yanagimoto (1992)、Kashiwagi (1993)）。また、世論調査結果の解析法として、多項系列データの平滑化（柏木・岸野 (1987)）や構造的変化の検知（Kashiwagi (1991)）などもある。

2. データと問題点

解析に使用するデータは、1975 年 1 月から 1980 年 6 月までの間に朝日新聞社、読売新聞社、毎日新聞社、共同通信社、および中央調査社の民間機関 5 社により実施された政党支持率調査の結果の内、自民党支持率の結果である。このデータを用いたのは、政党支持率は複数の機関により長期間継続的に調査されている極めて貴重な話題のひとつであり、また、期間中最も高い支持率を示したのが自民党であったためである。

データを図 1 に示す。横軸は年月、縦軸は支持率を表わし、各社の調査結果を A～E の記号により記してある。

図からすぐ分かるように、各社の調査結果には標本変動だけでは説明できない差異が存在す

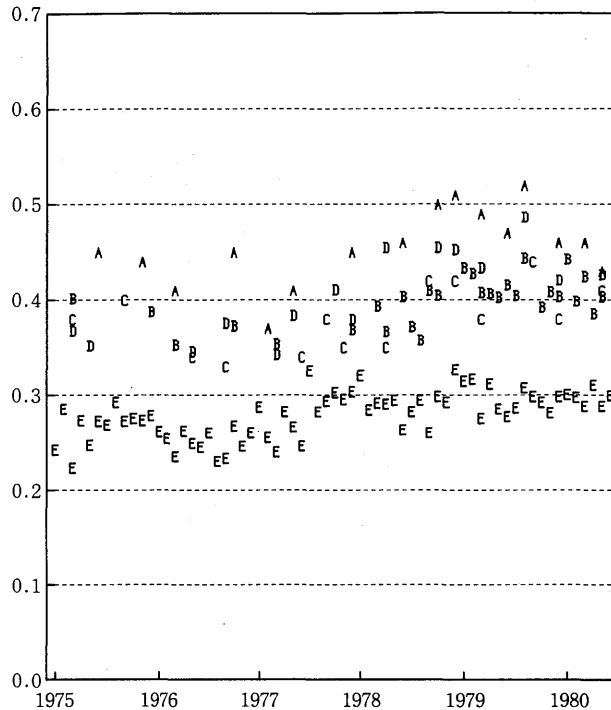


図1. 1975年1月から1980年6月までの間の朝日新聞社(A), 読売新聞社(B), 毎日新聞社(C), 共同通信社(D), および中央調査社(E)による自民党支持率。

る。一般に、これらの差異は調査技法の違いにより生じたと理解されている。実際、調査では、例えば、設問の違いにより調査結果に差異が生じるという事態は珍しくなく、調査技法は調査結果を左右する重要な要因となっている。ただし、調査結果を解釈する立場からすれば、そうした差異は望ましい存在ではない。なぜなら、無作為標本抽出を用い、同一の話題について調査しているにもかかわらず、調査技法の違いにより調査結果が異なれば、調査結果の客観的意味が不明となり、解釈が困難になるからである。実際、図1のように同一時点の支持率に0.2もの開きがあつては、支持率自体の解釈は難しい。

一方、図1のデータを継続調査の結果、すなわち、時系列として眺め、推移の様相のみに注目すると、各社の調査結果の差異は系統的であるようにも見える。つまり、各社の調査結果に共通のトレンドが存在し、それに系統的な差異を生じさせる成分が付加された可能性がある。共通トレンドが存在すれば、それは調査技法に依らない世論の動向を反映した客観的な情報といえる。なぜなら、各社の調査は各々独立に実施されているからである。そこで、共通トレンドの有無について検討する。

次章で、まず、共通トレンドが存在しない場合および存在する場合双方に対するモデルを仮定する。第4章で、これらのモデルに対する推定法について述べる。共通トレンドの有無は、仮定したモデルの内から最善のモデルを選択し判定する。モデル選択法については第5章で述べる。そして、第6章で、解析結果を示す。

3. モデル

本章では、解析に用いるモデルについて述べる。ここで、 y_{it} を i 番目の会社による t 時点における調査結果としておく。添字 i は図 1 中の A~E に対応し、 t は 1975 年 1 月から 1980 年 6 月までの各月に対応する。従って、今の場合、 $i=1, \dots, 5$ であり、 $1 \leq t \leq 66$ である。添字 t の範囲を不等号で表わしたのは、会社により欠測が存在するため、 t は各 i 毎に y_{it} が存在する時点全てを示すという意味で用いる。また、期間中における i 番目の会社の調査結果の数を n_i とし、 $n = \sum_i n_i$ とする。

まず、共通トレンドが存在しない場合のモデルとして、以下の各社毎に個別のトレンドが存在するモデルを仮定する。

$$\text{(Model Ia)} \quad y_{it} = x_{it} + u_{it}, \quad u_{it} \sim \text{i.i.d.} N(0, \sigma_i^2), \quad i=1, \dots, 5, \quad 1 \leq t \leq 66,$$

$$\text{(Model Ib)} \quad x_{it} = 2x_{i,t-1} - x_{i,t-2} + v_{it}, \quad v_{it} \sim \text{i.i.d.} N(0, \delta_i^2), \quad i=1, \dots, 5, \quad t=3, \dots, 66.$$

ただし、 x_{it} はトレンドであり、 u_{it} および v_{it} は残差である。Model Ia は観測の構造を表わし、Model Ib は「トレンドは滑らかに変化する」という先験情報を表わしている。以上の Model I は、標準的なベイズ型平滑化モデルである。なお、調査結果が支持率の形で与えられているにもかかわらず、残差 u_{it} の分布として上記の正規分布を仮定したのは、個々の調査の回収標本の大きさが不明であり、調査結果のレベルが各社毎に見ると比較的一定していたためである。

次に、共通トレンドが存在する場合について考える。共通トレンドが存在する場合のモデルとしては、様々なモデルが考えられるが、共通トレンドの存在を示すためには、Model I に対し何らかの意味において有意となる共通トレンドを含むモデルさえあればよい。精密化は次の段階の問題である。そこで、取りあえず、次の単純な 2 組のモデルを仮定する。

第 1 のモデルは、系統的な差異が加法的に生じるとする以下のモデルである。

$$\text{(Model IIa)} \quad y_{it} = b_i + x_t + u_{it}, \quad u_{it} \sim \text{i.i.d.} N(0, \sigma^2), \quad i=1, \dots, 5, \quad 1 \leq t \leq 66,$$

$$\text{(Model IIb)} \quad x_t = 2x_{t-1} - x_{t-2} + v_t, \quad v_t \sim \text{i.i.d.} N(0, \delta^2), \quad t=3, \dots, 66.$$

ただし、 x_t は t 時点における各社に共通のトレンドであり、 b_i は i 番目の会社の系統的な差を生じさせる成分で、識別可能性を考慮し $b_5 = 0$ とする。Model IIa は観測の構造を表わし、Model IIb は、Model Ib の場合と同様、「トレンドは滑らかに変化する」という先験情報を表わしている。以上の Model II は、分散分析における変量効果モデルの特殊な場合とも考えられる。

第 2 のモデルは、系統的な差異が乗法的に生じるとする以下のモデルである。

$$\text{(Model IIIa)} \quad y_{it} = b_i x_t + u_{it}, \quad u_{it} \sim \text{i.i.d.} N(0, \sigma^2), \quad i=1, \dots, 5, \quad 1 \leq t \leq 66,$$

$$\text{(Model IIIb)} \quad x_t = 2x_{t-1} - x_{t-2} + v_t, \quad v_t \sim \text{i.i.d.} N(0, \delta^2), \quad t=3, \dots, 66.$$

ただし、 x_t および b_i は Model II と同じ意味で用いるが、 b_5 に対する制約条件は $b_5 = 1$ とする。また、Model IIb と Model IIIb が同一である点を指摘しておく。

以上、3 組のモデルを仮定した。これらのモデルを比較し、Model II ないし III が最善のモデルとして選択されれば、共通トレンドの存在が示される。ちなみに、Model II が選択されると、モデルに基づく推定値をハット付きで書けば

$$\hat{y}_{is} - \hat{y}_{it} = \hat{b}_i + \hat{x}_s - \hat{b}_i - \hat{x}_t = \hat{x}_s - \hat{x}_t$$

となり、各社の観測変量の推定値間の差は、共通トレンドの推定値間の差に等しくなる。また、Model III が選択されると、

$$\hat{y}_{is} / \hat{y}_{it} = (\hat{b}_i \hat{x}_s) / (\hat{b}_i \hat{x}_t) = \hat{x}_s / \hat{x}_t$$

となり、各社の観測変量の推定値間の比は、共通トレンドの推定値間の比に等しくなる。従って、何れの場合にも、各社の観測変量の推定値からでも共通トレンドの推移に関する情報を抽出できる。一方、Model Iが選択されると、各社毎に個別のトレンドが存在するとしか言えず、調査技法に依らない情報を抽出するためには、モデルを更に工夫する必要が生じる。

4. 推 定

本章では、前章で仮定した3組のモデルに対する推定法について述べる。仮定したモデルは何れもベイズ型平滑化モデルであり、ベイズ型尤度と事後密度を用いて、未知変数および変量を推定する。詳細については Akaike (1980) を参照されたい。

4.1 Model I に対する推定法

推定法について述べるために、仮定したモデルを密度の形で書く。先ず、Model Ia であるが、これは次のように書ける。

$$p(\mathbf{y}_i | \mathbf{x}_i, \sigma_i^2) \propto \sigma_i^{-n_i} \exp \left\{ -\frac{1}{2\sigma_i^2} (\mathbf{y}_i - \mathbf{D}_i \mathbf{x}_i)' (\mathbf{y}_i - \mathbf{D}_i \mathbf{x}_i) \right\}, \quad i=1, \dots, 5.$$

ただし、 \mathbf{y}_i は i 番目の会社の調査結果 y_{it} を調査時点順に並べた $n_i \times 1$ のベクトルであり、 $\mathbf{x}_i = (x_{i1}, \dots, x_{i66})'$ である。また、 \mathbf{D}_i は y_{it} と x_{it} の添字を符合させるための 0 と 1 を要素とする $n_i \times 66$ の計画行列である。一方、Model Ib は次のように書ける。

$$p(\mathbf{x}_{si} | x_{i1}, x_{i2}, \delta_i^2) \propto \delta_i^{-64} \exp \left\{ -\frac{1}{2\delta_i^2} (\mathbf{E} \mathbf{x}_i)' (\mathbf{E} \mathbf{x}_i) \right\}, \quad i=1, \dots, 5.$$

ただし、 $\mathbf{x}_{si} = (x_{i3}, \dots, x_{i66})'$ であり、 \mathbf{E} は x_{it} の線形関係を記すための 0, 1 および -2 を要素とする 64×66 の計画行列である。Model Ib は

$$p(\mathbf{x}_i | \delta_i^2) \propto \delta_i^{-64} \exp \left\{ -\frac{1}{2\delta_i^2} (\mathbf{E} \mathbf{x}_i)' (\mathbf{E} \mathbf{x}_i) \right\}, \quad i=1, \dots, 5$$

として \mathbf{x}_i の improper prior としても扱えるが、本稿ではモデル選択をする必要があるため、 x_{i1} および x_{i2} を変量ではなく変数とみなし、前記のように \mathbf{x}_{si} の proper prior として扱う。

さて、推定であるが、Model I は会社毎に分離可能なので、以下のように会社毎に推定する。先ず、ベイズ型尤度

$$L_{Ii}(x_{i1}, x_{i2}, \sigma_i^2, \delta_i^2) = \int_{R(\mathbf{x}_{si})} p(\mathbf{y}_i | \mathbf{x}_i, \sigma_i^2) p(\mathbf{x}_{si} | x_{i1}, x_{i2}, \delta_i^2) d\mathbf{x}_{si}$$

を最大にする \hat{x}_{i1} , \hat{x}_{i2} , $\hat{\sigma}_i^2$ および $\hat{\delta}_i^2$ を求める。ただし、 $R(\mathbf{x}_{si})$ は \mathbf{x}_{si} の台を表わす。次に、これらの最尤推定値の下で、事後密度

$$p(\mathbf{x}_{si} | \mathbf{y}_i, x_{i1}, x_{i2}, \sigma_i^2, \delta_i^2) \propto p(\mathbf{y}_i | \mathbf{x}_i, \sigma_i^2) p(\mathbf{x}_{si} | x_{i1}, x_{i2}, \delta_i^2)$$

に関する \mathbf{x}_{si} の条件付き事後平均を求める。以上の手続きは、ベイズ型尤度と事後密度を具体的に書き下すと、結局、 \mathbf{x}_i に関する重み付き最小 2 乗問題と $\lambda_i = \sigma_i / \delta_i$ に関する単変数非線形最適化問題に帰着し、ハウスホルダー変換と直線探索法を用いて、容易に実行できる。モデル全体の尤度は

$$L_I = \prod_i L_{Ii}(x_{i1}, x_{i2}, \sigma_i^2, \delta_i^2)$$

により与えられる。

4.2 Model II に対する推定法

Model IIa は密度の形で次のように書ける。

$$p(\mathbf{y} | \mathbf{b}, \mathbf{x}, \sigma^2) \propto \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{Cb} - \mathbf{Dx})'(\mathbf{y} - \mathbf{Cb} - \mathbf{Dx}) \right\}.$$

ただし、 $\mathbf{y} = (y_1', \dots, y_5')$ 、 $\mathbf{b} = (b_1, \dots, b_4)'$ および $\mathbf{x} = (x_1, \dots, x_{66})'$ である。また、 \mathbf{C} および \mathbf{D} は y_{it} と b_i および x_t の添字を符合させるための 0 と 1 を要素とする $n \times 4$ および $n \times 66$ の計画行列である。一方、Model IIb は次のように書ける。

$$p(\mathbf{x}_s | x_1, x_2, \delta^2) \propto \delta^{-64} \exp \left\{ -\frac{1}{2\delta^2} (\mathbf{Ex})'(\mathbf{Ex}) \right\}.$$

ただし、 $\mathbf{x}_s = (x_3, \dots, x_{66})'$ である。

推定法は Model I の場合と同様である。先ず、ベイズ型尤度

$$L_{II}(\mathbf{b}, x_1, x_2, \sigma^2, \delta^2) = \int_{R(\mathbf{x}_s)} p(\mathbf{y} | \mathbf{b}, \mathbf{x}, \sigma^2) p(\mathbf{x}_s | x_1, x_2, \delta^2) d\mathbf{x}_s$$

を最大にする $\hat{\mathbf{b}}$ 、 \hat{x}_1 、 \hat{x}_2 、 $\hat{\sigma}^2$ および $\hat{\delta}^2$ を求め、次に、これらの最尤推定値の下で、事後密度

$$p(\mathbf{x}_s | \mathbf{y}, \mathbf{b}, x_1, x_2, \sigma^2, \delta^2) \propto p(\mathbf{y} | \mathbf{b}, \mathbf{x}, \sigma^2) p(\mathbf{x}_s | x_1, x_2, \delta^2)$$

に関する \mathbf{x}_s の条件付き事後平均を求める。以上の手続きも容易に実行できる。

4.3 Model III に対する推定法

Model IIIa は密度の形で次のように書ける。

$$p(\mathbf{y} | \mathbf{b}, \mathbf{x}, \sigma^2) \propto \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{Bx})'(\mathbf{y} - \mathbf{Bx}) \right\}.$$

ただし、 \mathbf{B} は 0 と b_i を要素とする $n \times 66$ の計画行列である。従って、Model IIb と同一である Model IIIb を併せて、ベイズ型尤度は

$$L_{III}(\mathbf{b}, x_1, x_2, \sigma^2, \delta^2) = \int_{R(\mathbf{x}_s)} p(\mathbf{y} | \mathbf{b}, \mathbf{x}, \sigma^2) p(\mathbf{x}_s | x_1, x_2, \delta^2) d\mathbf{x}_s$$

と書け、事後密度は

$$p(\mathbf{x}_s | \mathbf{y}, \mathbf{b}, x_1, x_2, \sigma^2, \delta^2) \propto p(\mathbf{y} | \mathbf{b}, \mathbf{x}, \sigma^2) p(\mathbf{x}_s | x_1, x_2, \delta^2)$$

と書ける。ただし、Model IIIa が非線形であるため、推定は Model I および II の場合ほど単純ではない。本稿では、以下の変形 EM-アルゴリズムにより推定する。

0. 初期値として、 $\hat{b}_i^{(0)} = 0$ および $\hat{x}_i^{(0)} = y_{5t}$ を与え、 $j=1$ とする。
1. E-ステップとして

$$\hat{b}_i^{(j)} = \frac{1}{n_i} \sum_t \frac{y_{it}}{\hat{x}_i^{(j-1)}}, \quad i=1, \dots, 5$$

を実行し、 b_i を推定する。ただし、 t は y_{it} が存在する範囲を動く。

2. $\max_{i=1, \dots, 5} |\hat{b}_i^{(j)} - \hat{b}_i^{(j-1)}| < \alpha$ ならば、 $j-1$ 回目の繰返しにおける最大尤度や推定値を最終的な値として、手続きを終了する。 α としては小さい数、例えば、0.001 を採る。終了しない場合は、 $\hat{b}_i^{(j)} \leftarrow 1$ とし、次のステップに進む。

3. M-ステップとして、 $b_i = \hat{\delta}_i^{(j)}$ という条件の下で、ベイズ型尤度および事後密度により b_i 以外の変数および変量を推定する。 b_i が固定されれば、Model IIIa は線形となるので、Model I および II の場合と同様の方法により推定できる。推定された各トレンドの値を $\hat{x}_i^{(j)}$ とし、 $j \leftarrow j+1$ としてステップ 1 に戻る。

5. モデル選択

仮定したモデルの内から最善のモデルを選択するために AIC (Akaike's Information Criterion, Akaike (1973)) を用いる。AIC は

$$AIC = -2 \times \log(\text{最大尤度}) + 2 \times (\text{自由変数の数})$$

により定義される。尤度としてベイズ型尤度を用いる場合には、AIC は ABIC (Akaike's Bayesian Information Criterion, Akaike (1980)) と呼ばれる。

さて、仮定した各モデルに対する AIC は、Model I, II および III に対するベイズ型尤度に含まれる自由変数の数が、各々、20, 8 および 8 であるため、

$$\begin{aligned} AIC_I &= -2 \log \prod_i L_{ii}(\hat{x}_{i1}, \hat{x}_{i2}, \hat{\sigma}_i^2, \hat{\delta}_i^2) + 40, \\ AIC_{II} &= -2 \log L_{II}(\hat{\mathbf{b}}, \hat{x}_1, \hat{x}_2, \hat{\sigma}^2, \hat{\delta}^2) + 16, \\ AIC_{III} &= -2 \log L_{III}(\hat{\mathbf{b}}, \hat{x}_1, \hat{x}_2, \hat{\sigma}^2, \hat{\delta}^2) + 16 \end{aligned}$$

により与えられる。これらを用い、AIC 最小を実現するモデルを最善と判断する。

6. 結 果

本章では、解析結果について述べる。

まず、図 2 に、Model I により推定されたトレンドを示す。会社 A および B のトレンドは湾曲した形状をしている。両社のトレンドの形状は、詳細には異なっているが、大局的には類似している。一方、会社 C および D のトレンドは直線的となった。また、会社 E のトレンドは直線的ながらもやや湾曲し、他 4 社の中間的な形状を示している。

次に、図 3 に、Model II により推定されたトレンドを示す。Model II による場合、各社のトレンドの形状は全く同一であり、図表的には共通トレンドが縦軸方向に移動しているにすぎない。にもかかわらず、その当てはまり具合は比較的良好といえる。ただし、子細に観察すると、例えば、会社 D のトレンドは前半が過大評価で後半が過小評価であるといった点も見られる。

更に、図 4 に、Model III により推定されたトレンドを示す。Model III の場合、各社のトレンドは共通トレンドの定数倍で与えられるから、形状は厳密には一致しないが、推移の様相は全く同一である。当てはまり具合を見ると、Model II の場合よりやや良好であり、Model II の会社 D に見られた偏りは僅かながら改善されている。

最後に、各モデルに対する最大対数尤度と AIC の値を表 1 に示す。AIC は Model III において最小となった。この結果により、Model III が最善と判断される。ちなみに、Model III は、自由変数の数が最小でありながら最大対数尤度が最大であるから、自由変数の数による補正抜きでも、最善と判断できる。同時に、Model I の自由変数の数を、例えば、 $\sigma_i^2 = \sigma^2$ および $\delta_i^2 = \delta^2$ として減らしても、Model I の AIC が最小となる可能性はないとも判断できる。

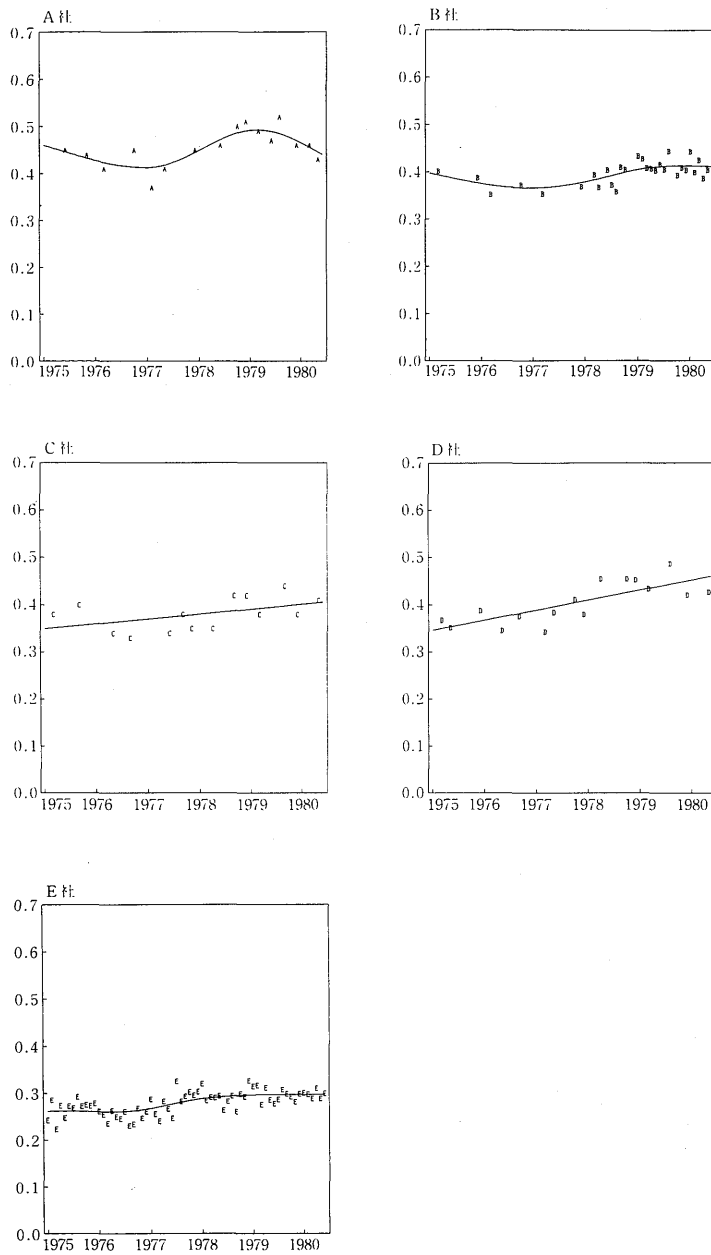


図2. Model I により推定されたトレンド。

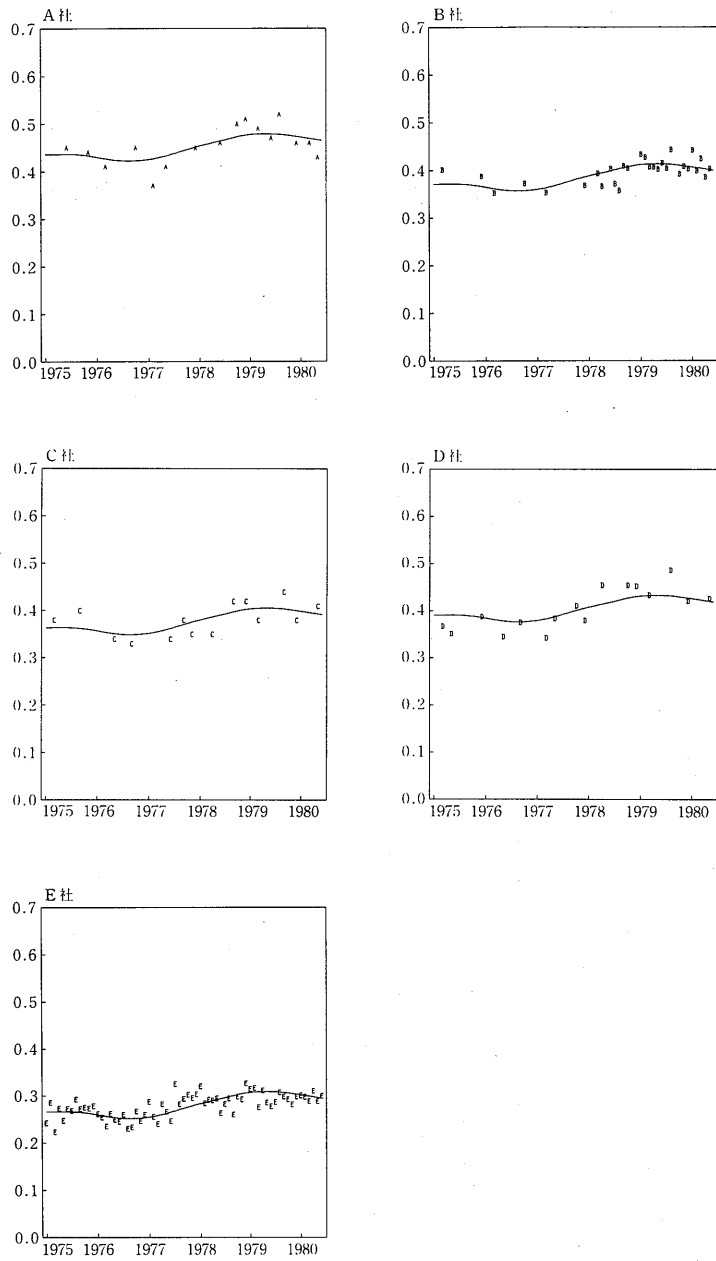


図3. Model II により推定されたトレンド。

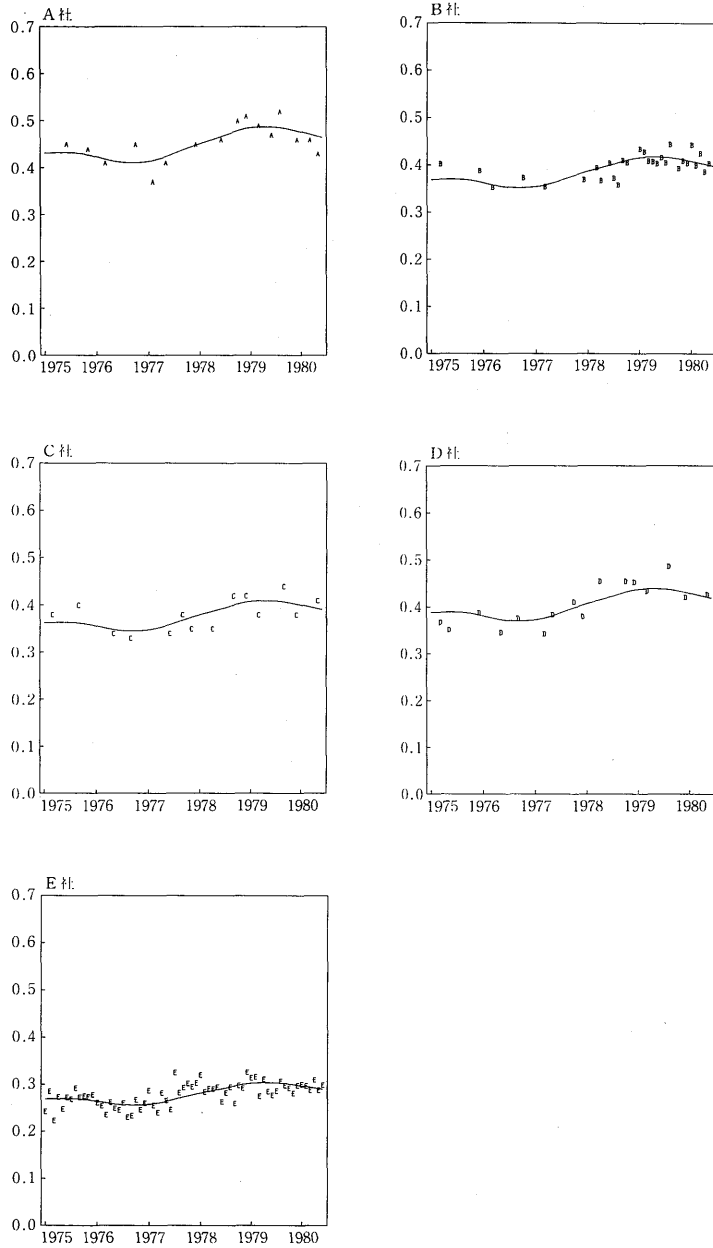


図 4. Model III により推定されたトレンド。

表1. 各モデルに対する最大対数尤度と AIC.

モデル	最大対数尤度	自由変数の数	AIC
I	-318.8	20	677.6
II	-317.6	8	651.2
III	-313.3	8	642.6

7. おわりに

調査技法の違いに起因すると考えられる差異を示す民間機関5社による継続世論調査の結果を、3組のベイズ型平滑化モデルを用いて解析した。その結果、各社の調査結果には調査技法に依らない共通トレンドが存在し、各社の観測変量の推定値からは共通トレンドの変化率を推定できるとの示唆が得られた。これは、世論調査の結果が調査技法に依存していても、継続調査と平滑化法により、調査結果から世論の動向に関する情報を抽出できる可能性を示している。

最後に、査読者の方々に感謝申し上げます。

参 考 文 献

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, *Second International Symposium on Information Theory* (eds. B.N. Petrov and F. Csaki), 267-281, Akademiai Kiado, Budapest.
- Akaike, H. (1980). Likelihood and the Bayes procedure, *Trabajos de Estadística*, **31**, 143-166.
- 柏木宣久 (1982). 圃場試験に於ける地力の推定, *統計研集報*, **30**, 1-10.
- Kashiwagi, N. (1991). Bayesian detection of structural changes, *Ann. Inst. Statist. Math.*, **43**, 77-93.
- Kashiwagi, N. (1993). On use of the Kalman filter for spatial smoothing, *Ann. Inst. Statist. Math.*, **45**, 21-34.
- 柏木宣久, 伊谷昭幸 (1986). 順序統計量とベイズ型平滑化条件に基づく条件付き百分位数の推定, *統計数理*, **34**, 213-220.
- 柏木宣久, 岸野洋久 (1987). 系列データの平滑化, *Information*, No. 8, 112-117.
- 柏木宣久, 柳本武美 (1990). 平滑化法とそのソフトウェア, *応用統計学*, **19**, 79-92.
- Kashiwagi, N. and Yanagimoto, T. (1992). Smoothing serial count data through a state-space model, *Biometrics*, **48**, 1187-1194.

Obtaining Information from Time Series Opinion Poll
Data Using a Smoothing Method
— A Case Study —

Nobuhisa Kashiwagi

(The Institute of Statistical Mathematics)

The usefulness of time series opinion poll data is shown through the analysis of actual data using a smoothing method. The data sets analyzed are the time series of the proportions of public support of the Liberal Democratic Party of Japan which had been collected by five companies independently from January 1975 to June 1980. Among the five companies' results, there exist some differences other than sampling variation. Such differences are believed to be caused by the differences among the survey techniques. To test whether the five companies' results involve objective information about public opinion in spite of the existence of such differences, three smoothing models are applied and the best fit one is selected among them using Akaike's Information Criterion. As a result of the analysis, it is suggested that, in the five companies' results, there exists a common trend which is independent of the survey techniques, and that the change rates of the common trend can be estimated from the estimates of the observed variables of each company. This shows the possibility that information about trend of public opinion can be obtained from time series opinion poll data using a smoothing method, even though the results of opinion polls depend on the survey techniques.