メディアにまたがる広告予算の配分における異なるラグ効果のインパクト

David Berkowitz
Arthur Allaway
Giles D’Souza

March/April 2001 Journal of Advertising Research

ビジネスにおける基本的な質問のひとつ、「広告にいくらかけるべきか？」それは主に、より基本的な質問への答えに依っている。「広告はどれだけセールスに影響するか？」プロモーション努力とセールス反応の関係のスタディはかように近代広告研究の基礎石となっている。1950 年代と 60 年代に、重要で多くの研究がセールス上のある期間とそれに続く複数期間における、広告の「残存」または「ラグ（遅れ）」効果に焦点を当て始めた。（例えば Jastram1955、Vidale と Wolfe1957、Palada1964、Tull1965 を見よ）本来、この残存効果における大きさと長さ、およびその推定は、広告予算の設定方法を改善する努力において重要なものである。

計量経済的モデルの文献として、広告とセールスの「遅れ」の関係を記した多くの研究が上梓されている。この分野における大部分の研究は、Koyck（1954）のラグ構造を採用しており、セールスにおける単一に集計された広告のラグ変数の効果を推定するアプローチをとる。実際、Leone（1995）の 30 年間の広告の残存に関するメタ分析は、複数のメディアによってひきおこされた別々のラグ効果の推定に関しては何も触れていない。複数メディアのラグ効果を扱った非常に少数のスタディはいずれも、Koyck のラグと 1 期間か 2 期間に限定された残存効果を先行研究とすることを選択している（Doyle と Saunders1990）か、またはこれらのスタディは暗黙的に異なるメディアは同じラグ効果を持つと仮定した推定テクニックを用い、わずか一つのラグ変数を推定している。

しかしながら、典型的な広告キャンペーンはいくつかのメディアが刺激しあって協調した様々な活動からなりたっている。もしあるメディアに特有のラグ構造が存在するならば（例えば、もしテレビ広告、プリント広告、看板広告、ラジオ広告がそれぞれ固有の残存効果を持つならば）、そのとき複数のラグ効果に基づいた広告予算配分プランは間違いなく一つのラグ効果に基づいたものとは違うだろう。このようなメディア固有の残存効果に関する研究は、広告支出額決定において、メディア間の予算*配分*は全体の予算*レベル*よりも重要かもしれないという最近の知見を得ていっそうふさわしいものとなっている。(Tull など 1986、Doyle と Saunders1990、Mantrala・Sinha・Zoltners1992、D’souza と Allaway1995)これより、我々は簡単な質問に答えようと思う。「もし異なるメディアに異なるラグ効果があるのならば、広告キャンペーンで使用される様々なメディアの長所をいかすように広告予算は配分されなければならないのだろうか？」

最初にまず、複数ラグに関する特別な考察を行うに当たって必要な論理的な基礎を探索

する。次に、広告予算配分におけるそれらの重要性を実証する。三番目に、複数ラグの推定方法を示す。最後に実際の広告データへの適用方法を述べる。

異なるラグの存在

　異なるメディアが同じラグ効果をもつべきであるとする強い理由はない。実際、直感的には異なるメディアは異なるラグ効果を持つはずである。さらには、ラグ効果の存在については 3 つの議論がある。最初は異なるメディアの認知的インパクト（強さ）に関する広告研究からのもので、2 番目は記憶の心理学スタディの成果から、三番目は受容された広告の結果と実践から得られたものである。幅広い広告研究の副題に見られる様々な発見が異なるメディアにおける異なるラグ効果の方向性を示してくれる。

異なるメディアの認知的強さ

　始めに、異なるメディアは異なる「現在」の効果をもつことは広告文献においては広く認められてる。((Doyle と Saunders1990）また「現在」の効果なくしては、長期的効果もありえない。(Jones1995）記憶の忘却のモデルを採用するにせよ、広告のラグ効果による記憶の干渉のモデル（Stewart1989）を採用するにせよ、広告のラグ効果は、現在効果との関連において異なるべきである。なぜならば、より強い記憶はよりゆっくりとした忘却を記録するだろうし、あるいはまた弱いものよりは干渉の危険がより少ないだろうからである。

　加えて、忘却率は素材が学習される度合いと、学習の集中度の両方に逆相関する。（Bean1912）異なるメディアにおいて、学習は異なる率で生起し、また忘却率は将来において変化する。これは中枢と末梢ルートの処理の概念（Petty・Cacioppo1981、1983）と一致する。Petty と Cacioppo の業績によると、態度変化は中枢処理ルートにおける勤勉な情報の考察の結果であり、長期保持が期待でき、比較的永続性を持つ。反対に、末梢ルートによる処理は刺激提示におけるほかの要因と関連付けられている。(環境や発話者など）そしてこのタイプの処理からおこる態度変化は一時的なようである。（Petty ・Cacioppo と Schumann1983）言葉によるあるいはビジュアルによる刺激はより中枢処理のトリガーになりやすいと示されてきており、(Beattie と Mitchell1985）こうしたメディアにおける効果は長期記憶により多くのインパクトを持ち、結果的にラグ効果を持つ。

　広告の反復に関する文献をレビューすると、Pechmann と Stewart（1989）は反復の結果におけるヴァリエーションのある一部は、露出の状況における差異によって明らかになると示した。メディアをまたがる露出の状況は定義によって異なるので、異なったメディアは異なった結果を生み、その結果違ったラグ効果を起こすと仮定するのが自然である。

　Chauduri と Buck（1995）は、電波メディアは印刷メディアより情緒的な面で関与を高

めると実証した。一方、プリントメディアはより論理的に関与を高めると示された。Lang と Dhillon と Dong（1995）は、TV コマーシャルにおいて、メッセージの arousal（刺激？）の度合い（情緒許容量？）と valence（原子価？）（肯定的または否定的尺度）が認知許容量と記憶に影響するとした。Silverman と Jaccard と Burke（1988）もまた、再生された属性は、すぐに再生された時期ではない、遅れて再生された時期において継続的であることを見出した。（意味不明。）これらの実践的要素はメディアによって変わるようなので、メディアごとに異なる現在のおよびラグの効果を期待することがもっともらしいと言える。

記憶の心理学的スタディ

異なったメディアに異なったラグ構造を想定する必要性を訴える心理学の文献からもまた証拠を見出すことができる。記憶研究において、リーセンシー効果（記憶リストにおける終了が終了前のｱｲﾃﾑよりすぐれた再生を示す）が視覚情報より聴覚情報において常に優れていると見出されてきた。この発見を外部情報という形式における干渉効果（Turner など 1995）の導入という考え方によってうまく説明する努力がなされているが、成功は一部にとどまっている。Sharps と Wilson-Leff と Price（1995）は記憶された関連する項目をｸﾞﾙｰﾌﾟ化することは（関連しないものをｸﾞﾙｰﾌﾟ化することに対して）絵画刺激より言葉刺激においてより高い効果が得られるとした。また、Assael（1992）は放送メディアはｲﾒｰｼﾞと象徴においてすぐれ、プリントメディアは細かい情報を伝えるのにすぐれていると断定した。異なるメディアはまったく固有の比率において様々な感覚刺激を結合しているので、異なるメディアが異なる記憶効果すなわち異なるラグ構造をもつと期待することがリーズナブルではないとはいえまい。

広告の結果と実務

最後に、広告の実務家は、異なるメディアは異なる結果を広告の露出後すぐに、またその後にも生むと絶対的に想定している。新聞、雑誌のような印刷媒体、TV やラジオのような放送媒体など、比較対象メディアにおいて CPM はかなり違うということを我々はよく知っている。広告の実務家はあるメディアはあることをするのに他のメディアより良いと信じているようである。（例えば、TV 広告はプリントよりインパクトがある、ラジオは柔軟性の点で最適なメディアである。）このようにメディアコストの多様性は異なるメディアが異なるレベルの時間的効果を持つということを多分に含んでいる。

複数ラグの重要性

研究者がセールスの反応において異なるメディアでの異なる効果の可能性を考えるべき

であるということはたやすい。異なるメディアにおけるラグ構造が同じであると考える理由はほとんどない。しかしながらもし異なるメディアが有意に異なるラグを持っていると示されたとしても、メディア間の広告予算配分においてこれらのラグインパクトがいかなる利益結果ももたらさないという議論が依然としてなされるであろう。何人かの研究者（Bultez と Naert1979 と 1988）は 1 変数ラグ構造による単純化は最適広告支出にほとんど影響がないと主張した。もしこれが複数ラグにおいても真実であり続けるのならば、彼らの経験的でエコノメトリックな研究はたいして重要ではないということになる。

一方、Magat・McCann と Morley（1986）はこの結論が維持されないような状況に焦点を当てたが、特にそれは広告が長期的ラグ構造を有する場合であった。この場合、広告は現在の期間よりもむしろ数期間前の方がより大きな効果を持つということになる。ブランド構築のための広告の場合のように、広告効果は長期間においてより大きなセールスへのインパクトを持つ。これは、消費者が評価したり回りの人たちと相談してから製品を買う（例えば自動車）時より起こりそうなことである。この結果利益の詳しい説明はますます困難になる？（Magat・McCann と Morley1986）

この問題を詳細に検討するために、我々はシミュレーション実験をセットし、遂行する。我々は二つの広告メディアのための簡単な加法広告反応モデルから始めることにする。

$$S(t)=\alpha+\sum_{i=0}^{i=\infty}\lambda_1^i f_1(A_1)+\sum_{i=0}^{i=\infty}\lambda_2^i f_2(A_2) \qquad (1)$$

ただし、$S(t)$ は t 期における広告収入を表し、方程式右辺の二項はある二つのメディア（たとえば新聞とビルボード）を代表する。これらの項において？、ある特定の期間における合計セールス $S(t)$ ＝*定数プラスメディア 1 における広告からの合計効果*プラス合計の*メディア 2 における広告からの合計効果*である。これらのメディアにおいて、関数 $f_1(A_1)$ と $f_2(A_2)$ は*現在*の広告のセールス収入に対するインパクトを特定し、λ_1 と λ_2 は*ラグ*のある広告効果を代表する、つまり過去の広告露出からの t 期におけるセールス収入への影響ということである。

広告予算決定時のジレンマは、二つのメディアにおける広告レベル A_1 と A_2 を、それぞれの単位コスト c_1 と c_2 と共に選ぶということであり、そうのような広告の貢献：

$$\Pi=\sum_{t=1}^{t=T}\{mS(t)-c_1A_1-c_2A_2\} \qquad (2)$$

は T 期間に渡って最大化される。m は貢献マージン率で Π は T 期間を通じたグロス広告貢献を表す。合計広告予算は X によって説明され、$X=c_1A_1+c_2A_2$ である。
シミュレーションは *3* タイプの広告予算モデル、*2* つの異なるコスト構造、*5* 個の現在効果パラメータ、*5* つの異なるまたあり得る、*それぞれの*メディアにおけるラグ構造、またパラメータ信頼に関する *5* つの異なるレベルでの不確定性を用いる。

広告予算モデルのタイプ

シミュレーションは広告予算配分の決定に通常用いられる 3 つのアプローチを通じて異なるラグ構造のインパクトをテストする。それらはカレント広告の、線形、定数弾力性、ADBUDG（Little, 1970）モデルである。図 1 は 3 モデルの広告のカレント効果とシミュレーションパラメータを示す。モデルは以下のように記述される。

1.線形 $f(A) = \beta A$

2.定数弾力性 $f(A) = \gamma A^{\beta}$

3.ADBUDG $(\gamma - \alpha) = \dfrac{A^{\beta}}{A^{\beta} + \delta}$

線形モデルは典型的に最もよく広告を説明し、定数弾力性と ADBUDG は共に比較的フラットである。

コスト構造

我々は二つのコスト状況をシミュレーションする：一つは各メディアのユニットコストが同じである場合で、もう一つは一方のメディアのコストは他メディアの 2 倍である場合である。最初の場合、$c_1 = c_2 = \$1$であり、2 番目のケースの場合、$c_1 = \$0.75$と$c_2 = 2 * c_1$である。

広告効果パラメータ

パラメータ値は 3 タイプのモデルを比較可能なように選ばれた。合計の広告ユニット$A_1 + A_2$が 1000 に等しい時、すべてのモデルは同じセールスレベルを与える。合計のラグ効果λがある。メディア 1 にラグ比率pがある（そして$1-p$がメディア 2)。メディア 1 にカレント効果比率qがある(そして$1-q$がメディア 2)。ラグ広告効果を説明するために、$\lambda_1 + \lambda_2 = \lambda$とし、$\lambda_1 = p\lambda$とする。ただし$0 < p < 1$。カレント広告効果を説明するために、最初のメディアに$f_1(A_1) = qf\ (A_1)$、2 番目のメディアに$f_2(A_2) = (1-q) f\ (A_2)$、ただし$q$は 0 から 1 の間に制限される。それぞれの可能性のある値λ、p、qを用いるよりも（これは潜在的な 1000 のコンビネーションを意味しただろう)、我々はλを 5 つの値に制限し（0.1、0.2、0.5、0.8、0.9)、pとqには独立に 5 つの値（0.05、0.3、0.5、0.7、0.95）を採用する。これらは統一されたレベルでのパフォーマンスを供給すべく選ばれた。それらはあらゆる可能性のある値の代表であるが、すべての値を含む必要がないということを意味している。

不確定性

最後に、我々はT（＝16）回繰り返される誤差項を式 1 に加え、5 レベルの不確定性を導入する。誤差項は平均 0 標準偏差σの正規分布と仮定され、異なるレベルの不確定性を代表するために 5 レベルに渡って変化する。最初のレベルで$\sigma = 0$完全確定のケースを代表

し、続く 4 つのレベルではσは 33 から 67、500、最後に 4500 へと増える。この方法では、不確定性のレベルは、完全確定性からランダム効果が平均的に決定的効果と同じぐらいであるような極端さにまで渡る。

テスト：カレント効果のみとカレント効果+ラグ効果の比較
広告ドルを配分する二つのアプローチはこのようにして比較された。相対的カレント効果のみによる配分方法は最初のメディアにおける$d = q$を設定するものである。（これを方法 1 と呼ぶ。）二番目の方法はカレント効果とラグ効果の双方を考慮に入れて配分方法を選ぶものである。この場合$d = \{q/(1-\lambda_1)\}/\{q/(1-\lambda_1)+\{(1-q)/(1-\lambda_2)\}$。（これを方法 2 と呼ぶ。）本質的に、広告効果が今期効果とラグ効果の率になるように選んだ。
　貢献マージン率mは 40 パーセントに固定され、シミュレーションは 500＄の初期値$S(0)$から開始された。2 つの方法はそのグロス広告貢献における効果を 3750 の観測値と比較された。

結果
　シミュレーション結果は 3750 のケースのうち、方法 2（メディアごとの今期効果とラグ効果）が方法 1（メディアごとの今期効果のみ）に勝ったケースが 2144 で、方法 1 が方法 2 より勝ったのが 1456 ケースであることを示している。150 ケースでは 2 つの方法に差がない。二つの方法に差がある時、方法 2 は 60%のケースでパフォーマンスが良い。これはその利益への貢献がこの 60%のケースで大きく、予算配分決定が、決定アプローチが今期とラグの効果の両方に責任を持つならばより良いということを意味している。
　表 1 にシミュレーションの結果をまとめた。方法 1 による予算設定は今期効果のみを考慮に入れている。方法 2 による予算設定は今期効果とラグ効果を考慮に入れている。（qは メディア 1 の今期効果の比率であり、pはメディア 1 のラグ効果の比率である。）
この結果は異なるメディアのラグ広告効果を無視することは重大な利益機会の損失を導くことがあることをも含んでいる。メディアにまたがって予算を配分するにあって、ラグとカレントの両方の効果を考慮に入れた方法（方法 2）はカレント効果のみを考慮に入れた方法（方法 1）よりも大多数のケースにおいてすぐれているようである。各メディアにおけるラグ効果の違い（広告ドルの残存力）が増し、メディアコストの違いも増すと、方法 2 の優位性も増す。図 2 はこの関係を示す。

推定
　シミュレーションの結果から、異なるラグ構造が重要な利益へのインプリケーションを持ち得、また更なる分析を保証すると言えるようである。次の問題はこれらを推定するための正しい方法は何かということである。（技術的な実証とどのようにモデルを推定したかについての説明はアペンディックスを参照のこと。）

データと結果

南東州における大きな西洋服販売の単独店が今回のスタディのためのデータを供給した。先行する二つのスタディが同じデータベースの一部にアクセスしており、一つは最適決定支援システムのデザインのためのものであり、(Allaway・Mason と Black1987) もう一つは広告支出決定のためのものである。(D'souza と Allaway1995) どちらのスタディも複数ラグを推定していない。データベースは 3 年間の、12 のデパートのデイリーの（ウイークリーに集計される）POS データと、同様に新聞広告のコラムのインチ数と、ある 1 時期に掲出されたビルボード（3 年間の間にはビルボードの数は 2 から 14 に動く）から構成されている。週ごとの新聞広告は主に全ページアピールに対する 1/8 ページ（メーカーによるクーポン広告）からなっており店舗寄りではなく製品寄りの訴求となっている。一方、ビルボード広告は、製品寄りではなくより店舗寄りとなっている。適当なチェックとクリーニングを経て、データは店舗セールスと新聞・ビルボード広告の 156 の週別観測値に集計された。

いくつかのバージョンの式（A6）がテストされ、結果的に以下のモデルが採用された。

$$(1-L\rho)S(t)=\beta_0+\beta_1A_1*(t)+\beta_2A_2*(t)+\beta_3t+\beta_4D_1+\beta_5D_2+\beta_6D_3+u(t) \quad (6)$$

ここで変数は以下を除き前出と同様である；

$A_1*(t)$ ：t 期のメディア i における好意または広告ストック

$i=1,2$ （1＝新聞広告のコラムインチ数、2＝ビルボードの数）

$t=1,2,....156weeks$

D_1 ：クリスマス時期だと値 1 を取りそれ以外だと 0 を取るダミー変数

D_2 ：学校に戻る時期だと値 1 を取りそれ以外だと 0 を取るダミー変数

D_3 ：製品ポリシーの変更が実施され、利益の少ない製品ラインが整理された後半の期間に値 1 を取りそれ以外だと 0 を取るダミー変数

ρ ：その他のマーケティング変数のラグ効果を反映するパラメータ、例えば店舗の位置、プロモーション、店舗内のサービス、ロコミなど。大部分のセールスは頻繁でない購入消費者の持続からきており（ウエスタンブーツとジーンズ)、平均パーチェスサイクルは 1 年を超えるので、ρ を購入強化と解釈するのはあやまりである。(Givon と Horsky1990)

モデルは二つの仮定をしている。最初に、同じメディアのすべての広告は同じ率で忘却する。二つ目に、ある特定のメディアを強化することことを助ける広告の間にメディアの忘却率の減衰を惹き起こすような交互作用は存在しない。例えば、セールス構築とブランド構築の効果は分けられるし、区別される。かように、我々のモデルはセールス構築のための広告の使用のみに焦点を当てておりブランドやブランド認知をサポートしていない。

SAS と EVIEWS という統計パッケージがモデル推定と診断チェックには使用された。モデル残差はそれらが古典的な回帰の仮定に従うかどうかを確かめるためにチェックされた。またモデルは係数の安定性をチェックするためにサンプルに渡って繰り返し推定された。

どちらのケースもモデルは適当であると認められた。

広告の好意ストックは新聞とビルボードに関し、λ_i を 0 から 1 に 0.05 ずつ増やした値として用い、A6 で定義されたように再帰的に構築された。これらの値の範囲は、各メディアにいかなるラグ効果も存在しないという仮説を含むことに留意しよう。（$\lambda_i = 0$）各メディアごとに 20 値の λ が好意比較のために使用されたわけで、20^2 あるいは 400 のコンビネーション、それはつまり式（6）が計算された回数である。

400 式のうちの 7 つと結びつけられた誤差平方和（SSE）を表 2 に示す。SSE が最小なのは $\lambda_1 = 0.20$ で $\lambda_2 = 0.35$ の時、最大なのは $\lambda_i = 0.95$ の時。明らかに有意なラグ効果がある。最も低い SSE を持つ 5 つの式はその λ_2 が 0.25 から 0.45 の間で、λ_1 は常に 0.20 である。

表 2 の示されたモデルの広告好意の今期効果に関しては表 3 で報告されている。これはモデル選択に当たっての厳密な最小 SSE 条件への固執の重要性をよく表している。さもなくば、混乱した豊富さの困惑に直面する。（？）自然に考えれば、現在のケースのように、最も少ない SSE を持ったモデル群が似通っているならば､すべてのモデルはよいといえる。しかし、もしモデルがまったく違うのならば、研究者はたぶん盲目的に最小 SSE ルールを採用することに気持ちの悪さを感じることになるだろう。その場合、選ばれたモデルはモデル間の差異を最大にしなければならない。

式 6 から推定されたパラメータは表 4 に示されている。Clarke は ρ を $0 < \rho < 1$、広告の持続間隔のパーセントを $\log(1-\rho)/\log(\lambda)$ として定義した。各メディアにこの式を使うと、我々の結果では 90％の継続間隔が新聞に関して 1.4 週であり（つまり $\log(0.1)/\log(0.2)$）、ビルボードに関して 2.2 週であり（つまり $\log(0.1)/\log(0.35)$）である。これはビルボードが新聞に比べて 53%長い（週の継続において）ということを意味している。また、回帰式のベータウエイトを見ると、ビルボードへ支出されたドルが追加の＄350.11 を創出し、新聞へ支出されたドルが追加の＄28.32 を生んだとされる。これは、この店とデータに関して、ビルボードがセールスに追加されるドルと広告の消費者の中に残る時間という両方の面においてより効果的であると提示している。

結論

我々は複数ラグ効果に関して 3 つの広い議論と､広告予算配分における重要性を実証し、またアクチュアルのデータからどのように推定したらよいかを示した。Ha（1995）はメディアモデルがリーチとフリークエンシーや露出確率に心を奪われてしまっており、それらの広告予算配分への使用はもっと改善できるとしている。メディアビークルのウエイト付けは質的量的両面の広告反応の違いを描画しようとする一つの試みである。にもかかわらず、独断による割り当てのために満足のいくものであるとはまるで言いがたい。経験的なあるいは歴史的でさえあるデータによって広告メディアの複数ラグ効果の推定を行うこと

は、独断的なメディア割り当てを廃止する確かな方法であると指摘するべきである。
実践的な見通しとして我々の結果はマネージャーが、特定のメディアにおける広告を同じドル結果を生むように調整すべく、いかにしてこの情報を使うかを描画している。例えば、その他のすべての定数をそのままにしてみれば、表 4 における式はビルボード広告はドルセールスへの貢献という意味において 12.36 倍、新聞より効果的であるといっている。これは、マネージャーは 1 回のビルボード広告からのセールス結果と同じにするためには新聞広告では 12.36 回打たなければならないといっているのだ。

制限と今後の研究

我々のスタディに適用されたいくつかの制限がある。最初に、複数ラグはカレント効果が 50 : 50 でラグ効果が 5 : 95 のときに適合したが、カレント効果が 50 : 50 である他のケースでは適合しなかった。これは我々のシミュレーションにおけるおぼつかなさであり、現実世界でこのようなことはありそうもない。次に、説明の簡潔性のために固定した予算を計算に使用した。これも現実世界ではありそうもない。追加研究によって異なるメディアの予算を変えることに焦点をあてることができるだろう。三番めに、我々のモデルは広告メディア間に交互作用が存在しないという盲目的な仮定を置いた。しかしながら、多くの「広告キャンペーン」が様々な形式でのメディアが使われることを強調または支援するようにデザインされている。特に、いくつかの広告はブランド構築のためであり、他の広告はセールスパフォーマンスを向上させるためのものである。加えて、広告はこれらのタイプのキャンペーンにおける忘却の理解に焦点をあてるべきである。4 番目に、我々の研究は一つのデータセットと一つの小売店から来たものである。そのため、ビルボードがこの市場でより効果的であると結論しているが、各市場は固有で効果（ラグとカレント）はそれぞれ推定されるべきである。最後にコースグリッドサーチは詳細をよりよいレベルにすることが可能である。しかしながら、我々の研究における主な強調点は、最もよいモデルを見つけるためのテクニックの開発である。将来の研究がよりよいグリッドとこれらを分けることが差異を生むかどうかの決定に焦点を当てることが可能である。

アペンディックス

我々の推定アプローチの実証において、我々はラグオペレータ L を用い、 $A(t)$ は t 期における広告効果を示すとした。ラグオペレータは 1 単位期間の従属変数の遅れ効果である。
すなわち、

$$LA(t) = A(t-1)$$

または、

$$(1-\lambda L)A(t) = A(t) - \lambda A(t-1)$$

また、

$$LLA(t) = L^2 A(t) = A(t-2)$$

広告のラグ構造は以下のようにラグオペレータを用いることで説明できる。

$$\frac{A(t)}{(1-\lambda L)} = A(t) + \lambda A(t-1) + \lambda^2 A(t-2) + \lambda^3 A(t-3) +,.... \qquad \text{(A1)}$$

ここで式 A1 において数学的に$|\lambda|$は＜1 であると定義される。直感的にλは、次の期に効果を持つt期の広告の比率を代表する。ラグ記号を用いると、Palda(1964)によってマーケティング文献に導入された、見慣れた Koyck モデルは以下のように書ける。

$$S(t) = \beta_0 + \frac{\beta_1}{(1-\lambda L)} A(t) + u(t) \qquad \text{(A2)}$$

または、項を整理すると、

$$(1-\lambda L)S(t) = (1-\lambda L)\beta_0 + \beta_1 A(t) + (1-\lambda L)u(t) \qquad \text{(A3)}$$

ここで、

β_0は切片

β_1はセールスへの広告の今期効果

λは広告のラグ効果で$0 \le \lambda \le 1$

$S(t)$はt期のセールス

$A(t)$はt期の広告

$u(t)$はt期のランダムな撹乱

t＝time

リサーチャーは式(A3)の推定に OLS（最小二乗法）を用い一般的に良い結果を得る。ただしこの場合 1 階の移動平均（MA(1)）による撹乱項を一般には無視する。

複数メディアのための Koyck ラグモデル

二つのメディアがそれぞれ独立したラグ構造をもつ場合の Koyck ラグモデルは以下のように表せる。

$$S(t) = \beta_0 + \frac{\beta_1}{(1-\lambda_1 L)} A_1(t) + \frac{\beta_2}{(1-\lambda_2 L)} A_2(t) + u(t) \qquad \text{(A4)}$$

記号は以下を除き前出と同様である；

β_iはメディア$i(i=1,2)$のセールスへの広告の今期効果

λ_iはメディア$i(i=1,2)$の広告のラグ効果で$0 \le \lambda \le 1$

$A_1(t)$はt期のメディア$i(i=1,2)$の広告

式の両辺を$(1-\lambda_1 L)(1-\lambda_2 L)$でかけて整理すると、

$$(1-\lambda_1 L)(1-\lambda_2 L)(S(t) - \beta_0 - u(t)) = (1-\lambda_1 L)\beta_1 A_1(t) + (1-\lambda_2 L)\beta_2 A_2(t) \qquad \text{(A5)}$$

ここで、二つのラグ入りのセールス項があり、一つは各メディアのラグ付き広告項でありもう一つは 2 階の(MA(2))移動平均撹乱項である。MA(2)の撹乱プロセスによる複雑さを無視して式（A5）に OLS を適用すると、7 つの係数の推定値が得られ、5 つのパラメータの

関数となっている。この複雑さゆえに、推薦される手続きは、$0 \le \lambda_1 \le 1$ と $0 \le \lambda_2 \le 1$ の間隔における 2 次元グリッドサーチであり、続いて一般最小二乗法である。Z と G の方法に従うと、推定式は以下のようになる。

$$S(t) = \beta_0 + \beta_1 A_1 * (t) + \gamma_1 \lambda_1^t + \beta_2 A_2 * (t) + \gamma_2 \lambda_2^t + u(t)$$

ここで $A_i(t)$ は再帰的に構成される。

$A_i * (1) = A_i(1)$

$A_i * (2) = (1-\lambda_i) A_i(2) + \lambda_i A_i * (1)$

……

$A_i * (t) = (1-\lambda_i) A_i(t) + \lambda_i A_i * (t-1)$

長い期間のデータであれば、λ_i^t の項は廃棄される。

訳者による注釈と解釈

Koyck ラグの式を今一度見よう。

$(1-\lambda L) S(t) = (1-\lambda L)\beta_0 + \beta_1 A(t) + (1-\lambda L) u(t)$

ラグオペレータはわかりにくいので、「マーケティングサイエンス」に従った展開式を示す。

$$S(t) = \beta_0 + \beta_1 (A(t) + \lambda A(t-1) + \lambda^2 A(t-2) + \lambda^3 A(t-3) + ...) + \varepsilon_t \qquad \text{2-1}$$

$$S(t) - \lambda S(t-1) = (\beta_0 - \lambda\beta_0) + \beta_1 A(t) + (\varepsilon_t - \lambda\varepsilon_{t-1}) \qquad \text{2-2}$$

または、

$$S(t) = (1-\lambda)\beta_0 + \lambda S(t-1) + \beta_1 A(t) + (\varepsilon_t - \lambda\varepsilon_{t-1}) \qquad \text{2-3}$$

これは、ε_t が互いに独立の正規分布に従うと仮定すると、$u_t = \varepsilon_t - \lambda\varepsilon_{t-1}$ は移動平均型の系列相関をする。Zeller、Giezel（1970）による推定法があるというが、見たことはない。

Koyck モデルと類似のモデルとして以下がある。

アドストックモデルではλの推定に 0 から 1 の間を細かく刻んでデータと適合度の高いλを選択するという方法が取られる。アドストックモデルは以下のように記述される。

$$S(t) = \beta_0 + \beta_1 At(\lambda) + \varepsilon$$
$$At(\lambda) = A(t) + \lambda A(t-1) + \lambda^2 A(t-2) + \lambda^3 A(t-3) + ...$$
2-4 と 2-5

このモデルは OLS を行うに何ら不都合はないが、λの推定はややあやふやな感じがある。

一方、ブランドロイヤリティモデルといわれるモデルは、2-3 式を直接定式化して、

$$S(t) = \beta_0 + \lambda S(t-1) + \beta_1 A(t) + \varepsilon_t \qquad \text{2-6}$$

とおく。このモデルは広告のラグを表現せず、この場合の λ は広告の残存ではなくブランドのロイヤリティを表す。

また当期効果自己回帰誤差モデルとの奇怪な名称のモデルでは、同じく広告の残存効果を仮定せずに、誤差項が自己回帰系列相関をする。つまり広告以外の要因が残存効果を持つものと仮定させる。式は以下のようになる。

$$S(t) = \beta_0 + \beta_1 A(t) + u_t \qquad 2\text{-}7$$

$$u_t = \lambda u_{t-1} + \varepsilon_t \qquad 2\text{-}8$$

2-8 に　$u_t = S(t) - \beta_0 - \beta_1 A(t)$ を代入し整理すると、

$$S(t) = (1-\lambda)\beta_0 + \lambda S(t-1) + \beta_1 A(t) - \lambda(\beta_1 A(t-1)) + \varepsilon_t \qquad 2\text{-}9$$

いずれのモデルも誤差項や従属変数の系列相関を無視すれば $S(t)$ に $S(t-1)$ や $A(t)$ あるいは $A(t-1)$ を回帰させて λ や β を推定できるように見えるが、かならずしもそうはいかないので少し厳密に考える。

「回帰分析とその方法」には以下のように各ケースの推定方法が簡単に触れられている。

従属変数に系列相関がある場合

$$S(t) = \beta_0 + \lambda S(t-1) + \beta_1 A(t) + \varepsilon_t \qquad 2\text{-}10$$

で表せる典型的な自己回帰型のモデルである。マーケティングモデルではブランドロイヤリティモデル 2-6 に相当する。ε_t は正規分布に従い、系列相関を持たないと仮定する。個の場合、従属変数 $S(t)$ は過去の値に依存し、各データが独立ではないので、OLS を行う場合に疑問が生じる。しかし、以下のように OLS が適当であることを示せる。

一連の A が与えられており、$S(1)$ も与えられているとする。この時観測値の集合は $\varepsilon_{2,}\varepsilon_{3,}\varepsilon_{4,}\varepsilon_{5,}......\varepsilon_n$ によって生成されるので、この観測値の集合を観測する尤度は、

$$p(\varepsilon_{2,}\varepsilon_{3,}\varepsilon_{4,}\varepsilon_{5,}......\varepsilon_n) = p(\varepsilon_2)p(\varepsilon_3).....p(\varepsilon_n)$$

である。正規分布を仮定しているので尤度関数を計算し、2-10 から、

$$\frac{1}{(\sqrt{2\pi\sigma^2})^{n-1}} e^{-(1/2\sigma^2)\sum(S(t)-\beta_0-\lambda S(t-1)-\hat{\beta}A(t))^2}$$

これは負の指数部分が最小のとき最大、すなわち、Min $\sum(S(t)-\beta_0-\lambda S(t-1)-\hat{\beta}A(t))$ の時 MLE（最尤推定）でこれは OLS である。

誤差項の系列相関

誤差に系列相関がある、すなわち以下の誤差項を仮定する場合。

$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$

ここで v_t は独立な正規分布に従う撹乱項。これは当期効果自己回帰誤差モデルの場合に相当する。

まず線形モデル、$S(t) = \beta_0 + \beta_1 A(t) + \varepsilon_t$ を仮定する。ρ を乗じた $t-1$ のものを引くと、

$S(t) - \rho S(t-1) = \beta_0(1-\rho) + \beta_1(A(t) - \rho A(t-1)) + (\varepsilon_t - \rho\varepsilon_{t-1})$

または

$S(t) - \rho S(t-1) = \beta_0(1-\rho) + \beta_1(A(t) - \rho A(t-1)) + v_t$

でこれは OLS に必要な条件を満たしており、1 階の差分による回帰である。ちなみにこの式は 2-9 における λ を ρ に変えたものに等しい。データの差分を取る場合、データが一つ減ってしまうので、それを防ぐために、最初の観測値の差分として以下の推定式がある。

$S(1) - \rho S(0) = \sqrt{1-\rho^2}\,S(1)$

$A(1) - \rho A(0) = \sqrt{1-\rho^2}\,A(1)$

ρ はこのままの OLS では推定できないので、

$\hat{\varepsilon}_t = \rho\hat{\varepsilon}_{t-1} + error$

に OLS を適用することによって推定する。これはつまり回帰式 $S(t) = \beta_0 + \beta_1 A(t) + \varepsilon_t$ によって得られた ε_t を ε_{t-1} で回帰することに他ならない。しかしここで得られる推定量 r は ρ を過小推定する可能性があるので、ダービンワトソンの検定を行う必要がある。仮に系列相関がこの検定により実証されたとしても、ρ の推定にはそのまま役立たないので、一つの案として以下がある。すなわち、式を 2-9 の形に戻し、$S(t-1)$ の係数（式 2-9 ではλ）を ρ の推定量とする方法である。

系列相関が誤差項と従属変数の両方にある場合

これは Koyck ラグの場合とよく似ている。つまり、

$S(t) = \beta_0 + \lambda S(t-1) + \beta_1 A(t) + \varepsilon_t$　　　3-1

$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$

のようなケースである。

ここで 3-1 式が $t-1$ でも成り立っていることに注目し、

$S(t-1) = \beta_0 + \lambda S(t-2) + \beta_1 A(t-1) + \varepsilon_{t-1}$　　　3-2

3-2 に ρ を乗じ 3-1 から引くと、

$S(t) = (\lambda+\rho)S(t-1) + (1-\rho)\beta_0 - \lambda\rho S(t-2) + \beta_1 A(t) - \rho\beta_1 A(t-1) + v_t$

ここで、

$v_t = \varepsilon_t - \rho\varepsilon_{t-1}$

なので望ましい性格を満足している。OLS を適用して、最後の二つの係数 β_1 と $-\rho\beta_1$ から

ρ を推定する。残りの方法は誤差系列相関の場合に等しい。
しかしこの方法では Koyck ラグの推定に役に立たないので、Koyck ラグの場合は前出の Zeller、Giezel（1970）による推定法を調べる必要がある。

ラグに関しては他にパスカルラグや Almon のラグ、あるいは時系列分析には ARIMA モデルなどがあるが力尽きたので別の機会にゆずる。