2012 年度日本計量生物学会年会講演予稿集 117-122
（2012 年 5 月 26 日，統計数理研究所）改訂版

経時測定データへの key-factor/key-stage 分析の応用

山村光司
国立研究開発法人農業・食品産業技術総合研究機構　農業環境変動研究センター
環境情報基盤研究領域　統計モデル解析ユニット

1.　はじめに

Key-factor/key-stage 分析は，もともとは生態学の分野で昆虫などの生命表を分析するために提案された分析法であり，ある発育ステージの個体数の決定に対して「どの要因（factor）がどの時点（stage）において効果を発現したか」について量的に示す記述的手法である (Yamamura, 1999)。しかし，この方法は医薬品の臨床実験において経時測定データを分析する際にも有効利用できる可能性がある (Yamamura, 2012)。本講演では，この分析法の概要を述べ，それがどのような形の示唆を与えてくれるのか，また，この分析を事前分析として用いることにより，どのような形のモデルを構築することができるのかについて議論する。

2.　Key-factor/key-stage 分析の概要

個体に何らかの実験処置をほどこした後に，いくつかの時点で値を測定するという場面を想定する。実験処理を要因（factor），測定時点をステージ（stage）とよぶことにする。総個体数を n，ステージ数を g とする。第 j ステージの測定値のベクトルを $\mathbf{y_j}$ とし（$j = 1, 2, \ldots, g$），便宜上 $\mathbf{y_0} = \mathbf{0}$ と定義すると，最終ステージにおける観測値 $\mathbf{y_g}$ を次のように表現することができる。

$$\mathbf{y_g} = \sum_{j=0}^{g-1} (\mathbf{y_{j+1}} - \mathbf{y_j}) \tag{1}$$

次に要因（factor）の影響について考える。要因の数を f とし，平均値と観測値との差 $\mathbf{y_j} - \bar{\mathbf{y}}_\mathbf{j}$ を次のように要因の線形式で表現することを考える。

$$\mathbf{y_j} - \bar{\mathbf{y}}_\mathbf{j} = \mathbf{X}\mathbf{b_j} + \mathbf{e_j} \tag{2}$$

ここに $\mathbf{X}$ はフルランクのデザイン行列である。$\mathbf{b_j}$ は第 j 番目のステージにおける係数を表す列ベクトルである。また，$\mathbf{e_j}$ は残差つまり個体差成分を示す列ベクトルであり，要因と共分散がゼロとなるように決められているとする。いま，$\mathbf{X}$ の中から h 番目の要因に関係する要素だけを残して他の要素をゼロに置き換えた行列を $\mathbf{X_h}$ とする（$h = 1, 2, \ldots, f$）。このとき2式は次のように書くことができる。

$$\mathbf{y_j} - \bar{\mathbf{y}}_\mathbf{j} = \sum_{h=1}^{f} \mathbf{X_h}\mathbf{b_j} + \mathbf{e_j} \tag{3}$$

すると，最終ステージの平均値まわりの平方和 $S(\mathbf{y_g})$ は次のように $(f+1) \times g$ 個の要素 r_{hj} に分割することができる。

$$r_{hj} = \begin{cases} (\mathbf{X_h}(\mathbf{b_{j+1}} - \mathbf{b_j}))'(\mathbf{X_h}\mathbf{b_g}), & j = 0,1,\ldots,(g-1), \quad h = 1,2,\ldots,f, \\ (\mathbf{e_{j+1}} - \mathbf{e_j})'\mathbf{e_g}, & j = 0,1,\ldots,(g-1) \end{cases} \tag{4}$$

それぞれの要素 r_{hj} は，第 h 番目の要因が第 j 番目のステージを通じて最終ステージの値($\mathbf{y_g}$) に

どれだけ影響を及ぼしているかを示しており，したがって r_{hj} を「最終結果に対する寄与（contribution）」と呼ぶことができる（詳細は後述の付録 D を参照）。これらの寄与を $(f+1)\times g$ の表に並べたものを key-factor/key-stage 分析表とよび, これらの寄与をグラフにして視覚的に示したものを key-factor/key-stage グラフとよぶことにする。

4式の各項は， $\mathbf{y_1}$ と $(\mathbf{y_{j+1}}-\mathbf{y_j})$ を多変量とみなして g 変量多変量分散分析を行う際に計算される平方和積和行列（sum of squares and cross products matrix; SSCPM）の要素と同一である。そのため，通常の統計解析ソフトウエアを用いることにより，直ちに key-factor/key-stage 分析を実行することができる。SAS 用のマクロと R 用の簡単な関数をそれぞれ付録 A および付録 B として次のサイトにおいた。http://cse.naro.affrc.go.jp/yamamura/Key-factor_analysis_program.html

もともとの key-factor/key-stage 分析では，要因には優劣関係が仮定され，指定された順番に要因が作用すると仮定されている（Type I 分割）。しかし，作用する順番に優劣関係がないと考えられる場合もあるであろう。そのような場合には，R 関数では「type=2」と指定し，SAS では「ss=ss2」と指定することにより，Type II 分割を行うことができる。この場合は，各要因に起因する変動と未知の変動を合計しても全体の変動には一致しないため，未知の変動はあまり意味を持たなくなる。そのため，Type II の key-factor 分析では Residual を除いた部分について寄与割合を計算する仕様にしてある。

また，上の計算では欠測値がない場面を想定しているが，実際にはそのような条件が満たされないことも多い。そのような場合には多重代入を行う必要がある。付録中の関数ではデフォルトでは多重代入はオフになっているが，これをオンに指定すれば，付録 A ではプロシジャ MI を用いて, また付録 B では MICE パッケージを用いて多重代入を行うようにしている (SAS Institute Inc., 2008; van Buuren and Groothuis-Oudshoorn, 2011)。また，反復数がそろっていない場合にはデフォルトでは Type I 分割を行うが，むしろ全観測値が欠測となっている仮想個体をデータ中に記入して，多重代入を行って計算を行う方が良いであろう。

3. 適用例１（１要因実験）

ある薬品会社は 3 種類の薬剤（標準薬，新薬，プラシーボ）がぜんそく患者の呼吸能力に与える効果を調べた (Littell et al., 2006)。各薬剤がそれぞれ 24 人の患者にランダムに割り当てられ，FEV1 と呼ばれる呼吸力を示す指数が 1 時間おきに 8 時間まで測定された。投薬に先立って 0 時点でも測定が行われている。FEV1 の患者毎の測定結果は図 1 に示してある。

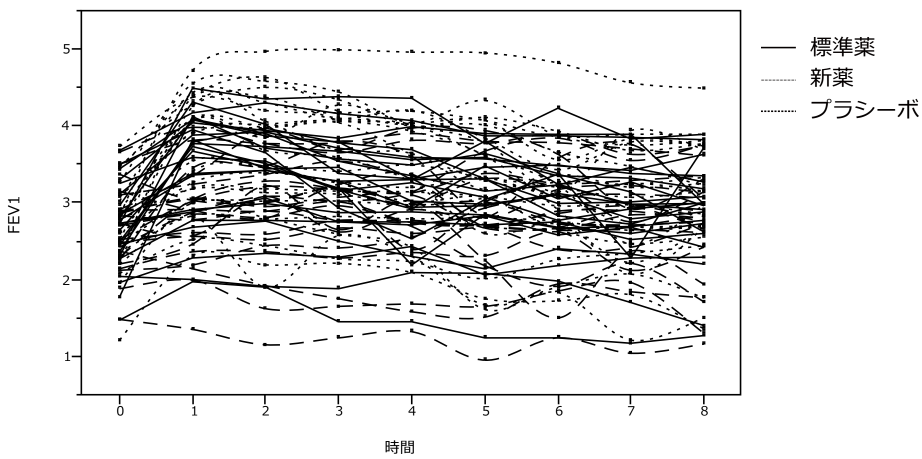


図 1. 各患者の呼吸能力 (FEV1) の時間変化。

Littell et al. (2006) はいくつかの議論の後に最終的に 7 パラメーターのモデルを採用した。

$$y_{vit} = b_0 y_{vi0} + b_{1v} + b_{2v} t + u_{vi} + e_{vit} \tag{5}$$

ここに y_{vit} は第 v 番目の薬剤の第 i 番目の患者の第 t 時点の FEV1 値である ($t = 0, 1, 2,..., 8$)。b_0 は薬剤投与直前の FEV1 に関する係数，b_{1v} と b_{2v} はそれぞれ第 v 番目の薬剤効果の切片と傾き，u_{vi} は平均ゼロの正規分布にしたがうランダム切片，e_{vit} は分散共分散構造 AR(1) にしたがう多変量正規分布誤差である。

この実験では要因は一つだけであるため，2式は次のように書ける。

$$y_{vij} - \overline{y}_j = \beta_{vj} + \varepsilon_{vij} \tag{6}$$

ここに y_{vij} は第 v 番目の薬剤の第 i 番目の患者の第 j ステージの FEV1 値である ($j = 1, 2,..., 9$)。$\overline{y}_j$ は第 j ステージの y_{vij} の平均，β_{vj} は第 v 番目の薬剤の第 j ステージにおける効果，ε_{vij} は β_{vj} に直交している残差である。混合モデルによる分析とは異なり，ここでは ε_{vij} について特別な分布や分散共分散構造を仮定しない。$t = 0$ 時点が $j = 1$ 番目の観測時点なので t と j は値が一つだけずれている。計算プログラム中では，たとえば以下のような形式で factor と stage を指定する。KFanalysis (factor = "factor(drug)", stage = "time0, time0to1, time1to2, time2to3, time3to4, time4to5, time5to6, time6to7, time7to8", data = asthma)。

図 2 に R プログラムから出力された key-factor/key-stage グラフを示す。このグラフ中の二つの線は 6 式の右辺の二つの成分に対応している。このグラフから得られる結論は以下のとおりである。結論１：薬剤効果は投与直後にだけ発生し（正の値），その後は薬剤効果は一貫して減衰する（負の値）。結論２：個人差は最初の 3 時間以内に発現し尽くして，3 時間より後には定常状態になる。結論３：個人差の大きさに比べると薬剤効果は小さく，標準薬と新薬，プラシーボ間の効果の違いは限定的である。

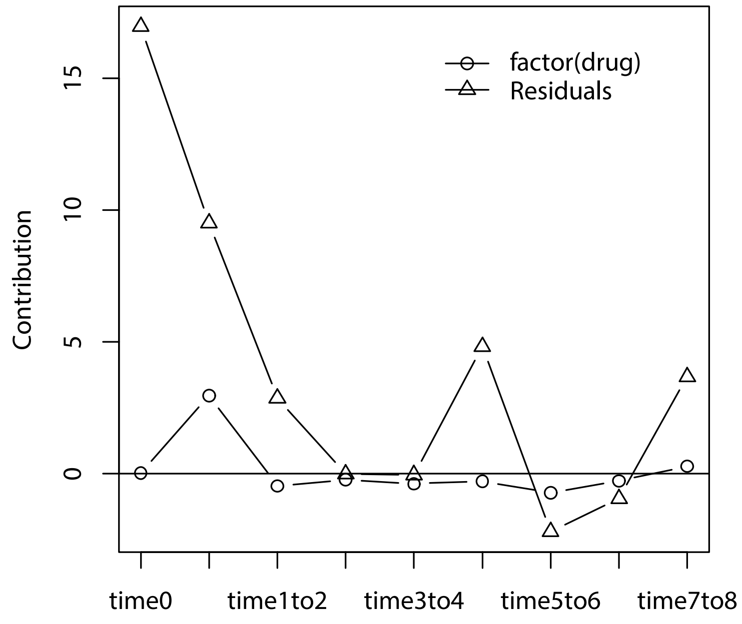


図 2. ぜんそく患者の呼吸能力改善に関する key-factor/key-stage グラフ。縦軸は r_{hj} を示す。付録 B の R 関数から描く。

予測を行う場面ではモデルの構築が必要となるが，その場合には key-factor/key-stage 分析を事前分析として用いれば有効なモデルが構築できる。図 2 を元にすれば次のようなモデルが構築できる。

$$y_{vit} = \begin{cases} b_0 + u_{vi} + e_{vi0} & \text{for } t = 0 \\ b_0 + u_{vi} + b_{1v}(1 - b_2(t-1)) + e_{vit} & \text{for } t \geq 1 \end{cases} \quad (7)$$

ここに e_{vit} は第 t 時点での正規分布にしたがう誤差成分である。図 2 より，0 時点での個体差成分が大きいことから，まずそれを $b_0 + u_{vi}$ として表現する。ここに b_0 は切片であり，u_{vi} は第 v 薬剤の第 i 個体の値である。図 2 は薬剤効果が投与直後にのみ生じることを示していることから，$t = 0$ と 1 の間の効果を b_{1v} によって表現する。さらに，図 2 は薬剤効果がそのあと一貫して減衰していることを示しているため, b_{1v} に減衰の傾き$(1 - b_2(t - 1))$ をかけている。7式の上式を下式に代入すれば次のような 4 パラメーターだけの式となる。

$$y_{vit} = y_{vi0} + b_{1v}(1 - b_2(t-1)) + (e_{vit} - e_{vi0}) \quad \text{for } t \geq 1 \quad (8)$$

Littell et al. (2006) をはじめ多くの人は, 薬剤投与前の測定値 y_{vi0} を共変量として扱っているが, 8式は, y_{vi0} を共変量として扱うべきではなく，その係数は常に 1 であるべきであることを示唆している。5 式と 8 式における誤差部分を便宜上いずれも CS+AR(1) で近似した場合には AIC はそれぞれ 198.2 と 194.2 であることから，8 式の予測力は 5 式の予測力よりも高いことがわかる。しかも，8 式は 7 式と同一の式であるから，そのパラメーターの生物的な意味は 5 式のそれよりも明快である (Yamamura, 2012)。

4. 適用例２（多要因実験）

Key-factor/key-stage 分析は，その本来の趣旨にあるように，多要因の観察や実験を扱う際にもっとも効果を発揮すると思われる。次に Stokes et al. (2001) が扱った多要因実験を分析する。この実験では患者は処理群とプラシーボ群にランダムに割り当てられ，そのあと 4 時点において呼吸状態が 5 段階で記録されている。試験は二つの医院（clinic）で行われた。その他の要因としては性別（sex），年齢（age）が記録されている。計算プログラム中では，たとえば以下のような形式で factor と stage を指定する。KFanalysis (factor = "factor(clinic), treatment, sex, age, factor(clinic):treatment", stage = "V0, V0toV1, V1toV2, V2toV3, V3toV4", data = resp)。

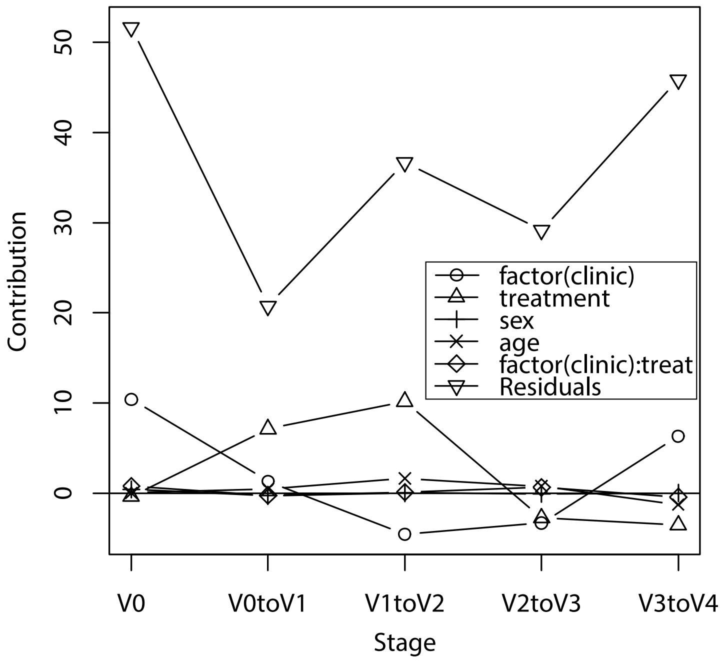


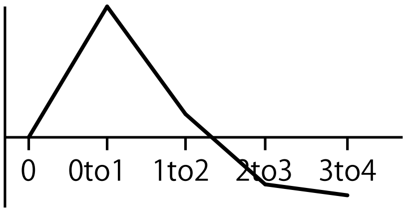
図 3．呼吸能力改善に関する Stokes et al. (2001) のデータについての key-factor/key-stage グラフ。縦軸は r_{hj} を示す。付録 B の R 関数から描く。

図 3 に R プログラムによって出力された key-factor/key-stage グラフを示す。このグラフから得られる結論は以下のとおりである。結論１：処理効果は最初の来院から 2 回目の来院までの間にだけ発現し，その後は減衰して処理前の状態に戻ってゆく。結論２：第 0 時点で clinic 効果が 0 でないことから, 初来院時には片方の医院に重篤な患者が集まっていたことがわかるが, その後に差は修正されている。結論３：図 2 と同様に個体差の成分はかなり大きいが，図 2 と違って個体差成分はゼロに収束しないことから，全期間を通じて個体差はランダムウォークのように集積されてゆくことがわかる。結論４：性別と年齢の影響はほとんど存在しない。

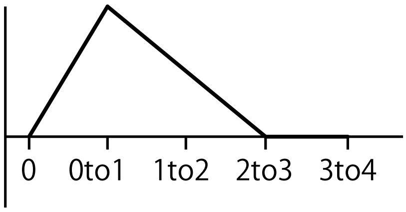
5. Key-factor/key-stage グラフの解釈

グラフを正しく解釈するには，少しばかり「慣れ」が必要である。要因効果成分と個体差成分について，一般に下図の例のように解釈ができる。なお，要因効果成分が実験開始前の時点（0 時点）でゼロから離れている場合には，要因効果がランダマイズされていないことを示唆している。

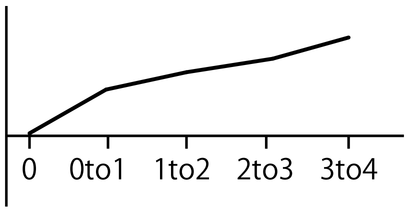
要因効果に関して



要因効果は 2 時点までに発現し，その後は効果は減衰して元の状態に戻ってゆく。

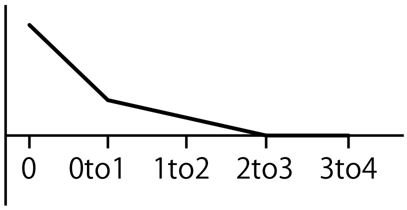


要因効果は 2 時点までに発現し，その後は効果は維持される。

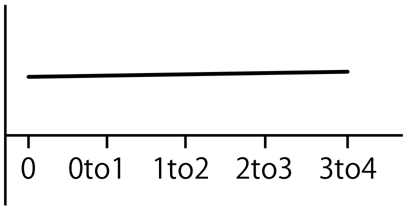


要因効果は加速的に大きくなる。（食事療法の効果など，効果の発現に時間のかかるデータで見られるパターン。）

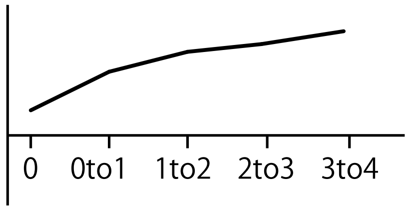
個体差 (Residuals) に関して



個体差は 2 時点までに発現しつくして定常状態になる。



個体差は集積されてゆく。（random walk かもしれない。）



個体差は加速度的に増加する。（random slope かもしれない。効果が相加的ではないので変数変換を考えるべきである。）

6. 文献

Littell, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. (2006). *SAS for Mixed Models*, 2nd edition. Cary: SAS Institute Inc.

Podoler, H. and Rogers, D. (1975). A new method for the identification of key factors from life-table data. *Journal of Animal Ecology* **44,** 85–114.

SAS Institute Inc. (2008). *SAS/STAT® 9.2 User's Guide*. Cary: SAS Institute Inc.

Smith, R.H. (1973). The analysis of intra-generation change in animal populations. *Journal of Animal Ecology* **42,** 611–622.

Stokes, M.E., Davis, C.S., and Koch, G.G. (2001). *Categorical Data Analysis Using the SAS System*, 2nd edition: Wiley-SAS.

van Buuren, S. and Groothuis-Oudshoorn, K. (2011). MICE:　Multivariate imputation by chained equations in R. *Journal of Statistical Software* **45,** 1–67.

Yamamura, K. (1999). Key-factor/key-stage analysis for life table data. *Ecology* **80,** 533–537.

Yamamura, K. (2012). Extended key-factor/key-stage analysis for longitudinal data. *Journal of Biopharmaceutical Statistics* **22,** 1–15.

連絡先：〒305-8604　茨城県つくば市観音台３－１－３，農研機構，農業環境変動研究センター，山村光司
E-mail: yamamura@affrc.go.jp，ホームページ: http://cse.naro.affrc.go.jp/yamamura/index.html

[付録 D] Key-factor/key-stage 分析の詳細

個体に何らかの実験処置をほどこした後に，いくつかの時点で値を測定するという場面を想定する。実験処理を要因（factor），測定時点をステージ（stage）とよぶことにする。総個体数を n，ステージ数を g とする。第 j ステージの測定値のベクトルを $\mathbf{y_j}$ とし，その平均値を並べたベクトルを $\bar{\mathbf{y}}_\mathbf{j}$ とする。つまり $\bar{\mathbf{y}}_\mathbf{j}$ は $n\times 1$ ベクトルであり，その要素はすべて y_j の平均値である。

最初のステージと最後のステージの観測値の差は次のように分解して表示することができる

$$\mathbf{y_g} - \mathbf{y_1} = \sum_{j=1}^{g-1} (\mathbf{y_{j+1}} - \mathbf{y_j}) \tag{9}$$

ここに $(\mathbf{y_{j+1}} - \mathbf{y_j})$ は第 j ステージから第 $(j+1)$ ステージへの変化を表す長さ n のベクトルである。$(\mathbf{y_{j+1}} - \mathbf{y_j})$ と $(\mathbf{y_g} - \mathbf{y_1})$ の平均値まわりの積和を k_j とする。つまり $k_j = [(\mathbf{y_{j+1}} - \mathbf{y_j}) - (\overline{\mathbf{y_{j+1}} - \mathbf{y_j}})]' \ [(\mathbf{y_g} - \mathbf{y_1}) - (\overline{\mathbf{y_g} - \mathbf{y_1}})]$。また，$(\mathbf{y_g} - \mathbf{y_1})$ の平均値まわりの平方和を $S(\mathbf{y_g} - \mathbf{y_1})$ とする。すなわち $S(\mathbf{y_g} - \mathbf{y_1}) = [(\mathbf{y_g} - \mathbf{y_1}) - (\overline{\mathbf{y_g} - \mathbf{y_1}})]'[(\mathbf{y_g} - \mathbf{y_1}) - (\overline{\mathbf{y_g} - \mathbf{y_1}})]$。このとき，9式を用いると $S(\mathbf{y_g} - \mathbf{y_1})$ を次のように分解することができる。

$$\begin{aligned}
S(\mathbf{y_g} - \mathbf{y_1}) &= [(\mathbf{y_g} - \mathbf{y_1}) - (\overline{\mathbf{y_g} - \mathbf{y_1}})]'[(\mathbf{y_g} - \mathbf{y_1}) - (\overline{\mathbf{y_g} - \mathbf{y_1}})] \\
&= [\sum_{j=1}^{g-1} (\mathbf{y_{j+1}} - \mathbf{y_j}) - \overline{\sum_{j=1}^{g-1} (\mathbf{y_{j+1}} - \mathbf{y_j})}]'[(\mathbf{y_g} - \mathbf{y_1}) - (\overline{\mathbf{y_g} - \mathbf{y_1}})] \\
&= \left(\sum_{j=1}^{g-1} [(\mathbf{y_{j+1}} - \mathbf{y_j}) - (\overline{\mathbf{y_{j+1}} - \mathbf{y_j}})] \right)' [(\mathbf{y_g} - \mathbf{y_1}) - (\overline{\mathbf{y_g} - \mathbf{y_1}})] \\
&= \sum_{j=1}^{g-1} k_j
\end{aligned} \tag{10}$$

したがって，これらの k_j を比較することにより，それぞれのステージにおける変化が全体の変化に対してどれだけの影響を持っているかについて評価することができる。この分散分割は Smith (1973) や Podoler and Rogers (1975) によって key-factor 分析法と呼ばれてきた。ただし，この分散分割では，全体の変化の変動を各ステージにおける変化の変動に分解しているため，key-factor 分析法というよりも key-stage 分析法を呼ぶべきであった。

最終ステージにおける観測値 $\mathbf{y_g}$ の平均値まわりの平方和も，10式と同様にして各ステージの要素に分割することができる。まず，便宜上 $\mathbf{y_0} = \mathbf{0}$ と定義する。すると9式から $\mathbf{y_g}$ を次のように表現することができる。

$$\mathbf{y_g} = \sum_{j=0}^{g-1} (\mathbf{y_{j+1}} - \mathbf{y_j}) \tag{11}$$

$(\mathbf{y_{j+1}} - \mathbf{y_j})$ と $\mathbf{y_g}$ の間の平均値まわりの積和 s_j とする。つまり $s_j = [(\mathbf{y_{j+1}} - \mathbf{y_j}) - (\overline{\mathbf{y_{j+1}} - \mathbf{y_j}})]'$ $(\mathbf{y_g} - \bar{\mathbf{y}}_\mathbf{g})$。ただし，$j = 1, 2, ..., g-1$ である。また，$\mathbf{y_1}$ と $\mathbf{y_g}$ の間の平均値まわりの積和を s_0 とする。つまり $s_0 = (\mathbf{y_1} - \bar{\mathbf{y}}_\mathbf{1})'(\mathbf{y_g} - \bar{\mathbf{y}}_\mathbf{g})$。さらに，$\mathbf{y_g}$ の平均値まわりの平方和を $S(\mathbf{y_g})$ とする。つまり，$S(\mathbf{y_g}) = (\mathbf{y_g} - \bar{\mathbf{y}}_\mathbf{g})'(\mathbf{y_g} - \bar{\mathbf{y}}_\mathbf{g})$。すると，11式によって $S(\mathbf{y_g})$ は s_0 と s_j の和として表現できる。

$$S(\mathbf{y_g}) = (\mathbf{y_g} - \bar{\mathbf{y}}_\mathbf{g})'(\mathbf{y_g} - \bar{\mathbf{y}}_\mathbf{g})$$

$$= \left(\sum_{j=0}^{g-1} (\mathbf{y_{j+1}} - \mathbf{y_j}) - \overline{\sum_{j=0}^{g-1} (\mathbf{y_{j+1}} - \mathbf{y_j})} \right)' (\mathbf{y_g} - \overline{\mathbf{y}}_\mathbf{g})$$

$$= \left(\sum_{j=0}^{g-1} [(\mathbf{y_{j+1}} - \mathbf{y_j}) - (\overline{\mathbf{y_{j+1}} - \mathbf{y_j}})] \right)' (\mathbf{y_g} - \overline{\mathbf{y}}_\mathbf{g})$$

$$= \sum_{j=0}^{g-1} s_j \tag{12}$$

次に要因（factor）の影響について考える。要因の数を f とし，第 i 要因の水準数を m_i とする。説明のため，要因はすべて名義要因であり，検査された個体数はどの要因の組み合わせについても同数（r 個体）である場合を最初に考える。このとき，検査個体数（n）は $n = r\Pi_{i=1}^{f} m_i$ として与えられる。平均値と観測値との差 $\mathbf{y_j} - \overline{\mathbf{y}}_\mathbf{j}$ を次のように要因の線形式で表現することを考える。

$$\mathbf{y_j} - \overline{\mathbf{y}}_\mathbf{j} = \mathbf{X}\mathbf{b_j} + \mathbf{e_j} \tag{13}$$

ここに $\mathbf{X}$ はフルランクのデザイン行列であり，いまは異なる要因のパラメーターは相互に直交している場面を考えている。$\mathbf{b_j}$ は第 j 番目のステージにおける係数を表す列ベクトルである。また，$\mathbf{e_j}$ は残差を示す列ベクトルであり，要因と共分散がゼロとなるように決められているとする。いま，$\mathbf{X}$ の中から h 番目の要因に関係する要素だけを残して他の要素をゼロに置き換えた行列を $\mathbf{X_h}$ とする。このとき13式は次のように書くことができる。

$$\mathbf{y_j} - \overline{\mathbf{y}}_\mathbf{j} = \sum_{h=1}^{f} \mathbf{X_h}\mathbf{b_j} + \mathbf{e_j} \tag{14}$$

すると，最終ステージの平均値まわりの平方和 $S(\mathbf{y_g})$ は次のように表すことができる。

$$S(\mathbf{y_g}) = (\mathbf{y_g} - \overline{\mathbf{y}}_\mathbf{g})'(\mathbf{y_g} - \overline{\mathbf{y}}_\mathbf{g})$$

$$= \left(\sum_{h=1}^{f} \mathbf{X_h}\mathbf{b_g} + \mathbf{e_g} \right)' \left(\sum_{h=1}^{f} \mathbf{X_h}\mathbf{b_g} + \mathbf{e_g} \right)$$

$$= \left(\sum_{h=1}^{f} \mathbf{X_h}\mathbf{b_g} \right)' \left(\sum_{h=1}^{f} \mathbf{X_h}\mathbf{b_g} \right) + \mathbf{e_g}'\mathbf{e_g}$$

$$= \sum_{h=1}^{f} (\mathbf{X_h}\mathbf{b_g})'(\mathbf{X_h}\mathbf{b_g}) + \mathbf{e_g}'\mathbf{e_g} \tag{15}$$

$S(\mathbf{y_g})$ は f 個の要因成分と残差成分に分けることができる。これは分散分析において昔から使用されてきた分散分割と同じである。15式の右辺の$(f+1)$個の項は，最終ステージの観測値 $\mathbf{y_g}$ に関する分散分析表の中の平方和の欄に書かれる値である。ただし，分散分析とは異なり，15式では残差成分についていかなる分布型もいかなる分散共分散行列も仮定されておらず，単なる記述を目的として分散分割が行われている。

12 式では最終ステージにおける分散 $S(\mathbf{y_g})$ を時間軸に対して縦方向（longitudinal）に分割しているのに対して，15式では $S(\mathbf{y_g})$ を時間軸に対して横方向（cross sectional）に分割している。この二つの分割法を組み合わせることにより，$(\mathbf{y_{j+1}} - \mathbf{y_j})$ と $\mathbf{y_g}$ の間の平均値まわりの積和 s_j を以下のように分解することができる。

$$s_j = \sum_{h=1}^{f} (\mathbf{X_h}(\mathbf{b_{j+1}} - \mathbf{b_j}))'(\mathbf{X_h}\mathbf{b_g}) + (\mathbf{e_{j+1}} - \mathbf{e_j})'\mathbf{e_g}, \qquad j = 0,1,\ldots,(g-1) \tag{16}$$

ただし$\mathbf{b_0} = \mathbf{e_0} = \mathbf{0}$である。16式と12式を用いると$S(\mathbf{y_g})$を$(f+1) \times g$個の要素$r_{hj}$に分割することができる。

$$r_{hj} = \begin{cases} (\mathbf{X_h}(\mathbf{b_{j+1}} - \mathbf{b_j}))'(\mathbf{X_h}\mathbf{b_g}), & j = 0,1,\ldots,(g-1), \quad h = 1,2,\ldots,f, \\ (\mathbf{e_{j+1}} - \mathbf{e_j})'\mathbf{e_g}, & j = 0,1,\ldots,(g-1) \end{cases} \tag{17}$$

それぞれの要素r_{hj}は，第h番目の要因が第j番目のステージを通じて最終ステージの値($\mathbf{y_g}$)にどれだけ影響を及ぼしているかを示している。これらの要素を$(f + 1) \times g$の表に並べたものをkey-factor/key-stage 分析表とよぶことができる。これらの要素をグラフにして視覚的に示すと，より効果的だと思われる。このグラフはkey-factor/key-stage グラフとよぶことができる。