

【研究紹介】

簡単にできると畜検査データの時系列分析

足立 泰基

北海道東藻琴食肉衛生検査所食肉検査課

(〒099-3231 網走郡大空町東藻琴千草72-1)

1. はじめに

多くの食肉衛生検査所は、衛生的な食肉生産を推進するために生産者にデータ還元を行っているが、統計処理まで行われている例は多くない。北海道東藻琴食肉衛生検査所では、と畜検査データを用いて時系列分析という方法で処理した結果を生産者に提供している[1]。と畜検査データの時系列分析について、近年いくつかの研究が発表されているが[2-4]、比較的高度な統計知識を必要とする。そこで、本稿は、時系列分析法を活用した、簡単にできると畜検査データの解析方法を紹介するとともに、種々の分野における時系列分析法を利用した研究を紹介する。

2. 時系列分析とはなにか

時間とともに変化するできごとを順番に測定することによって得られたデータを時系列データと呼ぶ。たとえば、毎月の平均気温、会社の四半期の売上高、毎晩測定した自分の体重といったものが時系列データにあたる。農場から搬入される牛や豚の毎月の臓器の廃棄数（廃棄率）も同様に時系列データである。

時系列データが完全にでたらめに変化するものではなく、なんらかの規則に基づいて変化するものであるならば、その規則性を推定することによって、変化の原因を考察して将来の変化を予測することや、逆に過去の予測を外れる変化を検出することによって、規則の変動が起こっている可能性を推論できる。このような、時系列データの規則性の推定を行うことや、推定された規則性に基づいて将来の予測を行うことなどを『時系列分析』と呼ぶ。

3. 時系列分析の例

(1) 牛乳の生産量に関する時系列分析の例

牛乳の生産量を毎月記録したものは時系列データであり、時系列分析の対象となり得る。この毎月の牛乳生産量の記録に規則性があるならば、数学モデルを作成することによって将来の予測を行うことができる。図1には、キューバの農業生産協同組合である共同生産基礎単位（UBPC）における牛乳生産量を実際に時系列分析した例を示す[5]。2010年までの時系列データをもとに、2011年の生産量を予測している。

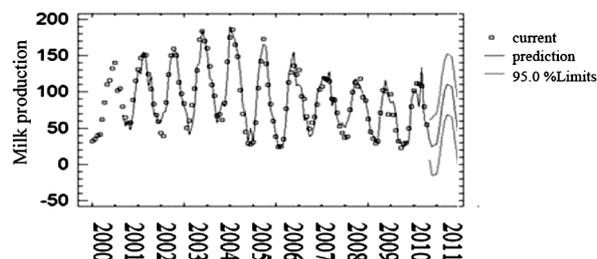


図1. 共同生産基礎単位（UBPC）マニアボにおける牛乳生産量の予測[5]。

(2) 牛の体温の時系列分析の例

牛の体温を一定時間ごとに継続的に測定したデータは、時系列データであり、時系列分析の対象となり得る。牛の体温は、発熱性疾患の発見以外に、ストレスの評価、分娩時期の予測や発情の発見などの指標として利用価値が高い。実際に牛に生体情報ロガーを取り付けて体温をモニターし、時系列分析することにより異常体温を検出した研究の例を図2に示す[6]。

図2の○印は、コンピュータによって検出された異常

連絡責任者：足立 泰基 北海道東藻琴食肉衛生検査所食肉検査課

〒099-3231 網走郡大空町東藻琴千草72-1

TEL：0152-66-2001 FAX：0152-66-3576 E-mail：adachi.yasumoto@pref.hokkaido.lg.jp

体温で、異常体温検出の判別のために、過去12日間の同一時間帯の平均気温との差からスチューデントのt値を求めるという方法が取られている。

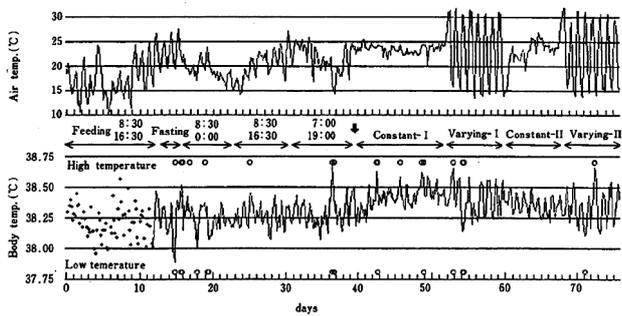


図2. コンピュータによる異常体温の判別[6].

上グラフは、周辺の気温、下グラフは測定された体温である。○：コンピュータによって判別された異常体温・・・：分析準備期間の測定値↓：環境調節室への入室時点

(3) 感染症患者数に関する時系列分析の例

ブラジルのリオデジャネイロ市内における毎月のデング熱患者発生数は時系列データであり、時系列分析の対象となり得る。実際に、リオデジャネイロ市内における毎月のデング熱患者発生数(1997年1月～2004年12月)の規則性を時系列分析によって見出し、2005年の患者数推移を予測した研究の例を図3に示す[7]。図3の縦のラインより右側に表示されているのが予測値のグラフで、上限の折れ線は95%信頼限界を示している。

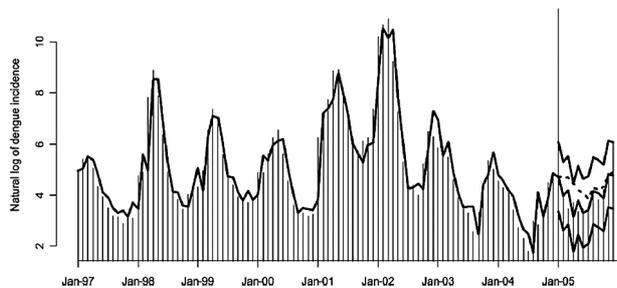


図3. リオデジャネイロにおけるデング熱患者発生数の推移[7].

縦軸は自然対数スケール

(4) と畜検査データの時系列分析の例

食肉衛生検査所による毎日のと畜検査で抗酸菌症と判定された豚の数は、時系列データであり、時系列分析の対象となり得る。実際に、と畜検査で抗酸菌症と判定された豚の数(2003年4月～2011年3月)の規則性を時系列分析によって見出し、抗酸菌症と判定される豚の数を予測(2011年4月～2013年5月)し、95%信頼限界を示している研究例を図4に示す[4]。図4において点線で示されているのが95%信頼限界で、95%信頼限界を超える

判定数が見られた日の搬入業者の中に、アウトブレイク農場が含まれていることを疑うのに利用できる。

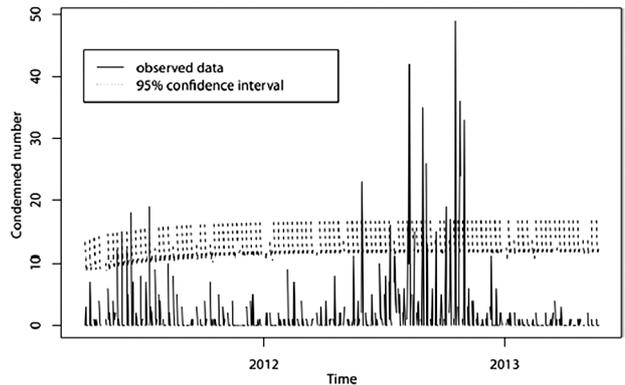


図4. 抗酸菌症と判定された豚の数(実線)と、時系列モデルによって計算された期待値の95%信頼限界(点線)[4]

4. と畜検査データの時系列分析にトライしてみよう

(1) 指数平滑法[8]

時系列分析には多様な方法があるが、多くは難解で煩雑である。ここでは、簡単な処理でモデルを作成することができる指数平滑法について紹介する。指数平滑法は過去の値に基づき、未来の予測値を計算する時系列分析法の一つで、「前の期の誤差に一定のウェイト(α)をかけて、その値に1つ前の予測値を加えたものを次の期の新しい予測値とする」という方法である。式で記載すると次のようになる。

$$y_{t+1} = \alpha(x_t - y_t) + y_t$$

ここで、 x_t は t 期における実測値、 y_t は t 期における予測値、 α は過去の値に掛けるウェイトで平滑定数と呼ばれるものである。

(2) ソフトウェアのダウンロードとインストール

分析用のソフトウェアとして、Rを用いる場合を説明する。Rは、オークランド大学のIhakaとGentlemanによって作られた統計解析用プログラミング言語で、多くの研究・教育機関で用いられている。高い機能を持ち、無料で自由に用いることができる。最新のRを以下のサイトからダウンロードし、インストールすることができる。

<https://cran.r-project.org/>

(3) 指数平滑法による時系列分析の例

ここでは、モデルケースとして毎月200頭搬入している農場xがあり、x農場の毎月の寄生虫性肝炎の廃棄数

	A
1	38
2	28
3	20
4	26
5	14
...	...
77	88
78	63
79	61
80	80
81	

図5. 過去データファイル“xfarm.csv”の内容

の時系列分析を行ってモデルを作成し、未来1年間の廃棄数の予測を行う場合を考える（単純なケースを考えるため、搬入頭数は毎月同じ場合を想定）。廃棄数は、図5のようにエクセルのcsvファイルに一行に過去のデータから順に記録されており、xfarm.csvという名前で保存されているとする。

Rを起動し、「ファイル」メニューから「ディレクトリの変更」を選び、“xfarm.csv”が存在する場所を指定する。以下のようにコマンドを入力すると、“xfarm.csv”の内容がdataという名前の変数（オブジェクトと呼びます）に読み込まれる。左端の>はRがコマンド入力を求める記号である。

```
> data<-read.csv (“xfarm.csv”, header=FALSE)
```

dataは、ただの数値であるため、以下のコマンドで年月がわかる時系列オブジェクトに変換する。frequency=12で1年=12カ月周期の時系列オブジェクトに変換される（frequency=4とすれば四半期となる）。

start=c(2003, 4)は、2003年4月が最初のデータであることを示している。

```
> tsdata<-ts (data, frequency=12, start=c(2003, 4))
```

tsdataと入力すると中身を確認することができる。

```
> tsdata
```

```
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
2003      38 28 20 26 14 18 28 22 18
2004 23 15 9 24 21 15 14 33 20 14 31 24
2005 11 18 16 13 12 33 19 19 . . . .
```

次に、HoltWinters関数を用いて指数平滑法の平滑定数を推定し、時系列モデルを作成する。推定結果をfitというオブジェクトに収容する。fitと入力することにより、推定結果を見ることもできるが、ここでは省略する。

```
> fit<-HoltWinters (tsdata, beta=FALSE, gamma=FALSE)
```

作成した時系列モデルを用いて最終データ（ここでは、

2009年11月）の翌月から1年間の予測値と信頼限界を計算する。予測にはforecast関数がよく用いられるが、標準関数の中に含まれていないので、CRANのミラーサイトからダウンロードする必要がある。パッケージメニューから「CRANミラーサイトの設定」（日本国内のサイトを通常選択する）を行い、「パッケージのインストール」から、forecastを選択してダウンロードする。インストール済みであれば、以下のコマンドでforecastライブラリを呼び出した後、forecast関数を実行する。

```
>library (forecast)
```

```
>yosoku<- forecast (tsdata, h=12)
```

h=12は12カ月分の予測を意味する。数字を変えることにより予測期間を変えることができる。

結果をグラフに表示するには、plotコマンドを使用する。以下のように入力すると観測値と1年間の予測値を表示される。グラフを図6に示す。

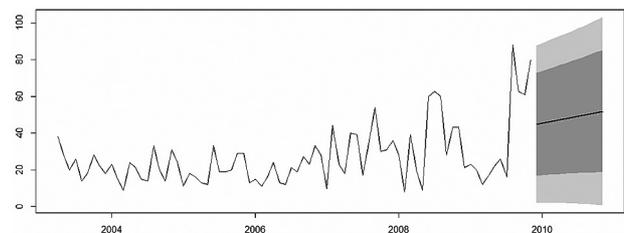


図6. plotコマンドによって出力された寄生虫肝炎による肝臓廃棄数の推移と、予測された80%ならびに95%信頼限界。

```
>plot (yosoku)
```

グラフの縦軸は頭数である。2009年12月以降に表示されたバンドが予測の範囲となる。濃い部分が80%信頼限界の領域で、薄い部分を含めた全体が95%信頼限界である。levelオプションを用いて、これらの範囲を変えることができる。たとえば、50%と80%を表示するのであれば、以下のコマンドを入力する。

```
>yosoku<-forecast (tsdata, level=c(50, 80), h=12)
```

図6と同様に以下のplotコマンドによって出力した例を図7に示す。

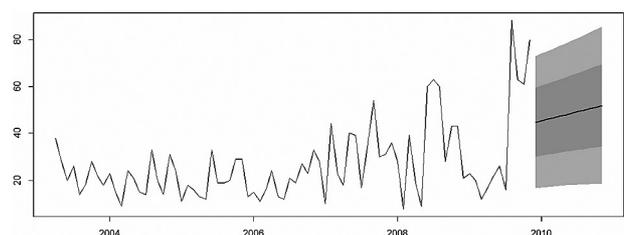


図7. 図6の信頼限界を50%と80%に変更したグラフ。

```
>plot (yosoku)
```

以下のコマンドをさらに入力することにより、モデルによって計算された期待値が点線で表示される。期待値の表示を加えたグラフを図8に示す。

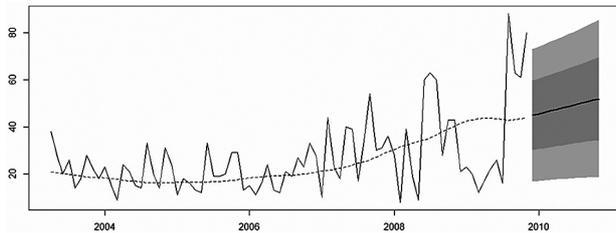


図8. モデルによって計算された期待値(点線)を図7に追加表示したグラフ。

```
>lines (yosoku$fitted, lty=2)
```

以上、と畜検査データを用いて初歩的な時系列分析を行う方法の概略を説明した。この例では、Rの持つ高度な機能のごく一部のみが用いられている。時系列分析の方法は多数あり、Rで利用できる方法も多い。いろいろな方法にチャレンジしていただき、と畜検査データの新たな活用方法を開拓するきっかけとしていただければ幸いである。

なお、北海道東藻琴食肉衛生検査所のホームページには、データ還元のページ (http://www.okhotsk.pref.hokkaido.lg.jp/ds/hse/new_HP/datakangen.htm) を設けており、時系列分析による統計処理済み還元データの例とグラフの読み方を説明したファイルをダウンロードできるので、時系列分析を利用したデータ還元を行う上で参考にして頂きたい。

引用文献

- [1] 足立泰基、蒔田浩平：季節自己回帰和分移動平均モデルによると畜検査データの時系列分析法、日獣会誌、68、189-197 (2015)
- [2] Vial F, Reist M: Evaluation of Swiss slaughterhouse data for integration in a syndromic surveillance system, BMC Vet Res, 10, 33 (2014)
- [3] Neumann E, Hall W, Stevenson M, Morris R, Ling Min Than J: Descriptive and temporal analysis of post-mortem lesions recorded in slaughtered pigs in New Zealand from 2000 to 2010, N Z Vet J, 62, 110-116 (2014)
- [4] Adachi Y, Makita K: Real time detection of farm-level swine mycobacteriosis outbreak using time series modeling of the number of condemned intestines in abattoirs, J Vet Med Sci, 77, 1129-1136 (2015)
- [5] Sánchez L, Cabanas G, Abad Y, Torres V: Use of ARIMA models for predicting milk production. Case study in UBPC "Maniabo", Las Tunas, Cuban, J Agr Sci, 48, 213-218 (2014)
- [6] 柏村文郎、新出陽三：生体情報ロガーを用いた牛の体温測定と時系列分析、日畜会報、61、869-875(1989)
- [7] Luz PM, Mendes BV, Codeço CT, Struchiner CJ, Galvani AP: Time series analysis of dengue incidence in Rio de Janeiro, Brazil, Am J Trop Med Hyg, 79, 933-939 (2008)
- [8] Holt C: Forecasting seasonals and trends by exponentially weighted averages, Int J Forecast, 20, 5-13 (2004)