

ジェボンスの「太陽黒点説」の再検討

西尾 敦
中野 聡子

1. はじめに

M.S. Morgan はその著書⁽¹⁾の第 1 部において景気変動分析の歴史を解説し、冒頭の節を1875年に発表された William Stanley Jevons (1835–1882) の太陽黒点説⁽²⁾にあて、これを黎明期の研究として取り上げている。

Jevons は18世紀後半から19世紀のイギリスでみられたほぼ10年を周期として繰り返される恐慌 (crisis) の原因を自然現象に求め、ほぼ同じ周期⁽³⁾を持つ太陽活動が、天候に影響を与え、さらに農業生産、農産物の価格へと順にその影響が及び最終的に経済活動全般の周期性を生み出すと信じた。Jevons は、このアイデアを証明するべく、データを探し求めたが、思うようなものが得られなかった。彼は、それは複雑化した近代社会では農業生産は様々な要因の影響を受けるからであると考え、古い時代では経済が単純であるから、より直接的に太陽活動の影響を受けるであろうとの理由で中世に注目し、Rogers によって編集された⁽⁴⁾イギリスにおける13–14世紀の各種穀物価格の膨大なデータ⁽⁵⁾を用いて、農業生産に11年周期があることを証明しようとした (論文VI)。しかしながら、Jevons は後に、この論文で用いた「方法」で、3年、5年、7年、9年さらには13年の周期性も11年周期と同様に「Rogers データ」にあてはまった、としてこの論文 (論文VI) を取り下げってしまった⁽⁶⁾。Jevons は、論文VIの後にも「太陽黒点説」にこだわり続けて、2編の論文を書いた⁽⁷⁾。

太陽黒点と経済活動の周期性の間に何らかの関連を見いだそうとする試み、すなわち「太陽黒点説」は、Jevons が経済学あるいは経済統計学においての他の業績によってすでに名声を獲得

(1) Morgan, M.S. *The History of Econometric Ideas*, Cambridge, Cambridge Univ. Press, 1990

(2) The Solar Period and the Price of Corn, 1875, Foxwell, H.S. *Investigation in Currency and Finance*, Macmillan, London, 1884, 2nd.Ed. 1909に論文 VI pp.175–186 として収録、以後「論文VI」と呼ぶ。

(3) 現代では、ほぼ11.1年とされている。

(4) Rogers, J.E.T. *A History of Agriculture and Prices in England*, Oxford, Clarendon Press, 1866

(5) 以後、「Rogers データ」と呼ぶ。

(6) Jevons, *Commercial Crises and Sun-Spots*, Part I, 1878, 前述 Foxwell 1909, に論文VIII pp.201–221 として収録、以後「論文VIII」。

(7) Jevons, *The Periodicity of Commercial Crises and Its Physical Explanation*, 1878, 前述 Foxwell 1909, に論文VII pp.187–200 として収録、以後「論文VII」。および論文VIII。

していたにもかかわらず、同時代の研究者の間で荒唐無稽であるとして概して不評であった。また、その後の統計学、計量経済学の歴史の中でも、太陽黒点説はあまり注目されることがなく、とりわけ論文VIは「著者も認めた誤りである」として「片付け」られているようである⁽⁸⁾。

Jevons が「誤った」理由は、偶然変動に起因する「見かけの事実すなわち周期性」を発見してしまったことである。このことは、Jevons にすくなくとも論文VIにおいて「確率論的視点」すなわち「有意性」の概念がなかったことが原因であると思われる。

現代の統計学には、データの整理・記述あるいはモデルによるデータの背景の記述という側面と、理論仮説のデータに基づく検証の2つの側面がある。後者を一言に要約すれば、データが示す結果が偶然変動によるものか否かを確率を用いて判断することである。現代の統計的検定の原型を、不十分ながら確立したのは Edgeworth である⁽⁹⁾。それに10年先だって発表された論文で Jevons が犯した誤りは「教訓的」である。彼の時代から1世紀余りを経た現代の（初歩的な）方法と道具（コンピュータ）を用いて再解析を行うと、その解析のどこが問題であり、どの点が評価できるかがあぶり出されてくる。ひいては、なぜ統計的検定が必要であるか、またモデルすなわちデータの得られる背景に依存して様々な方法が必要となるのかが明らかになる。本稿では、「Rogers データ」の現代的な解析例を示し、Jevons の解析を評価することを目的とする。また、その過程で、適切な統計解析が必要とされる理由を初学者が理解するための好例が、副産物として得られたと思う。

2. 季節変動の分析

Jevons は太陽黒点説に取りかかる前に、経済活動の周期的変動の解析に関して、すでに実績があった⁽¹⁰⁾。そこでの「周期変動」は、カレンダーに依存する制度（あるいは気候）に起因する「季節変動」が中心であった。この場合、周期は「年」あるいは「四半期」であり、データは月次あるいは四半期を最小単位として得られる。Jevons は「季節変動」のパターンが、その他の不規則な「変動」によって覆われ、不明瞭になっていると考え、「サイクル内の同じ位置」（周期が年であれば同じ月あるいは四半期）ごとの平均を取ることによって、不規則変動を消し去ることができると考えた。現代風に表現すれば、観測系列 $x_t, t = 1, \dots$ のモデルは

$$x_t = \mu + S_{m(t)} + e_t \quad (1)$$

(8) 例えば、J.Klein は *Statistical Visions in Time Series—A History of Time Series Analysis 1662–1938*, Cambridge, Cambridge Univ.Press, 1995. のなかで Jevons の研究に多くのページを割いているが、おそらく Jevons 自身が取り下げたという理由であろうか、論文VIについては全く言及がない。

(9) Edgeworth, F.Y. *Methods of Statistics, Jubilee Volume of the Statistical Society*, pp.181–217, 1885

(10) On the Study of Periodic Fluctuations Commercial, 1862, 前述 Foxwell 1909, 論文 I pp.1–12. On the Frequent Autumnal Pressure in the Money Market and the Action of the Bank of England, 1866, 前述 Foxwell 1909 論文 V pp.151–174. など。

と表される。ここで、 t は形式的な時刻（たとえば「月」）の通し番号、 $m(t)$ は t に対応する「月」の番号、 e_t は不規則変動を表す項、 μ は全体平均であり S_j , $j = 1, \dots, p$, は季節成分であり、 $\sum_{j=1}^p S_j = 0$ と仮定する。また p は周期であり月次データの場合 $p = 12$ 。現代の季節変動モデルでは、 $\mu = \mu_t$ は時刻 t と共に緩やかに変動する趨勢変動成分である。なお、時系列が各成分に「分解される」と仮定し、これらを移動平均によって「推定」することは Jevons の時代にはまだ明瞭に意識されていない。(1)式で表されるモデルが正しければ、 $m(t)$ の値ごとにデータをグループ⁽¹¹⁾化し、各グループごとの平均によって S_j を推定することができる（ e_t をほぼ消し去ることができる（大数法則））。Jevons は、このようにして（ $\mu + S_m(t)$ によって表される）季節変動パターンを抽出した⁽¹²⁾。

ところで、(1)のようなモデルの妥当性は、次のように検証できる（よく知られた 1 要因モデルの分散分析法）。日本の名目 GDP 系列（2006–2010 年、四半期系列）を例にとる。

表 1 GDP

	Qtr1	Qtr2	Qtr3	Qtr4
2004	121.3	124.1	122.1	130.9
2005	121.5	125.2	122.8	132.2
2006	122.9	126.3	123.9	134.2
2007	126.5	128.5	125.5	135.1
2008	126.6	126.2	122.2	130.0
平均	123.8	126.1	123.3	132.5

図 1 GDP

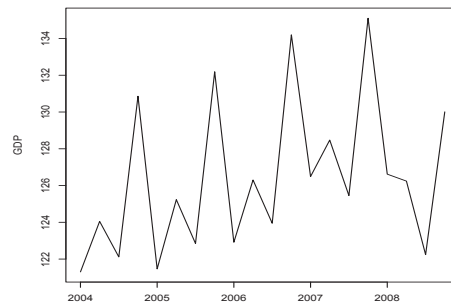


表 1 は、各行が 1 周期すなわち各列がそれぞれ第 1 四半期、第 2 四半期、...に対応するようにデータを配置した表である。これをチャート化した図1 では、同じパターンが繰り返されている様子が明瞭である。（(1)で誤差項に独立同一な正規分布を仮定する）素朴なモデルでは、分散分析の結果は、表 2 にまとめられる⁽¹³⁾。

季節指数 S_j , $j = 1, 2, 3, 4$ の有意性（すなわちグループ平均の差が偶然ではないこと）の程度は、表中の F に対応する p - 値によって判断される。この場合、 p - 値は 6×10^{-6} であるから高度に有意である。

さて Jevons は、その論文において月平均のみを記している⁽¹⁴⁾。論文の記述だけからは誤差分散を知ることができず「季節平均」の有意性は判定できない。しかしながら、季節変動は一般に経済データにおいて顕著であり、多くの場合上の例のように高度に有意になるので、Jevons はおそらくは有意な季節性を導き出しているものと推察できる。

(11) 本稿では、カタカナ化した外来語について、ラ行で表記される音のうち原語で “r” に相当するものには濁点をつけ “l” の音と区別した。

(12) たとえば、論文 I の plate I, p.11。

(13) 10節、参照。

(14) 論文 I, pp.9-10, 論文 V, p.163 など。

表2 分散分析表 (GDP データ)

要因	変動	自由度	不偏分散	F	p-値
季節性	268.31	3	89.44	22.11	6×10^{-6}
誤差	64.73	16	4.05		
全体	333.04	19			

3. 「Rogers データ」の11年周期性の分散分析

Jevons は、太陽活動の農業生産への影響を傍証するため、「Rogers データ」に見られる（はずの）11年周期性を立証しようとした（論文VI）。データは、小麦（wheat）、大麦（barley）、オート麦（oats）、豆（beans）、エンドウ豆（peas）、カラスノエンドウ（vetches）、ラ"イ麦（rye）の7種⁽¹⁵⁾についての1259年から1400年までの価格系列である。Jevons は、各系列から1314年の値を取り除いた上で、1259年の価格、1270年の価格、...と続く第1グループ（サイクル第1年）、1260年の価格、1271年の価格、...と続く第2グループ（サイクル第2年）、...、というように、「サイクル年」ごとの11グループに分け、季節性の分析と同じようにグループごとの平均値を求めた。表3は、Jevons の集計結果である。

表3 「Rogers データ」の周期性 (Jevons)

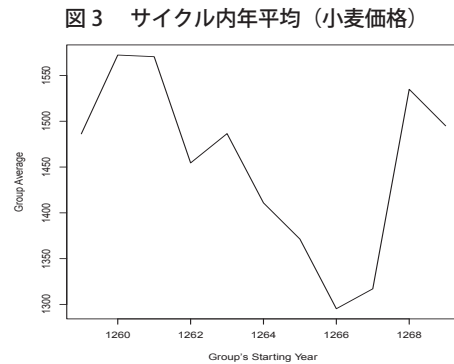
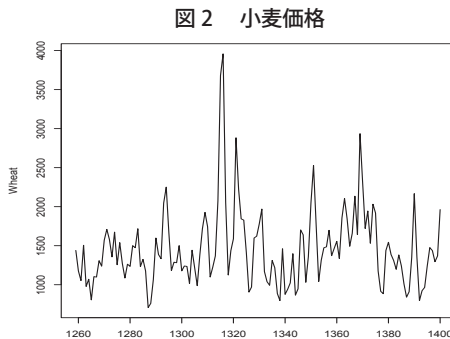
Years	1st 1259 etc.	2nd 1260 etc.	3rd 1261 etc.	4th 1262 etc.	5th 1263 etc.	6th 1264 etc.	7th 1265 etc.	8th 1266 etc.	9th 1267 etc.	10th 1268 etc.	11th 1269 etc.
Wheat	1490	1490	1570 (m.x)	1570 (m.x)	1460	1490	1410	1370	1290 (m.n)	1320	1540
Barley	1058	1021	1026	1111 (m.x)	979	1017	1010	1027	965	902 (m.n)	1036
Oats	598	596	619	699 (m.x)	593	625	592	575	581	555 (m.n)	626
Beans	993	1064	1243	1424 (m.x)	1111	1116	980	1095	983	843 (m.n)	1049
Peas	928	818	1019	1231 (m.x)	1016	891	870	880	757 (m.n)	767	929
Vetches	900	848	1139	1228 (m.x)	1016	953	916	824	850	920 (m.n)	808
Rye	1047	1025	1245 (m.x)	1161	1011	1142	1065	1139	1027	968 (m.n)	1094
Sums	7014	6862	7861	8424	7186	7234	6843	6910	6453	6275	7082

(m.x) denotes that the preceding number is the greatest in the line, and (m.n) similarly the least.

ここでは、小麦価格系列を取り上げて、季節性の例にならって1要因モデルを考え、Jevons が行わなかった方法（分散分析法）で11年周期性の検証を行ってみよう。

(15) 前出 Rogers の vol.1 pp.236-44. 原資料にはこの7種の他に drage（日本名不明）、2種のモルト（麦芽）の合計10系列が記されている。このうちモルトについては欠測が多いため分析が困難であったものと考えられるが、drage の欠測の状態は採用された系列と同じ程度である。これらを除いた理由は論文に記述されていない。

原系列をプロットすると、図2を得る。これには、図1のような明瞭な周期性は観察されない。しかし、サイクル内年のグループごとに平均を求めプロットすると、図3のようになる。これらの図を一瞥すると11年を周期とするゆっくりとした変動があるように思われなくてもいい。



定量的にこれを検証するために、分散分析を行ってみる⁽¹⁶⁾。各サイクル年ごとのグループの繰り返し数を揃えるために、データは1259年から1390年までの132年分を用いた。結果は表4のとおりである。このケースで $F=0.667$ に対する p -値は75%である。図3に見られる「サイクル内変動」は全く偶然であるといっても良い。

表4 分散分析表（小麦価格）

要因	変動 ×1/1000	自由度	不偏分散 ×1/1000	F	p -値
周期変動	174	10	17.4	0.667	0.75
誤差	3162	121	26.1		
全体	3336	131			

4. 2 要因モデルの分散分析—Jevons の表

前項の解析で、Jevons が主張する周期性は小麦価格系列において有意ではないと結論した。

(16) ここでは素朴な分散分析を行うが、このデータには変数間の相関および時系列的な系列相関が存在し、誤差の独立性を仮定する通常の分散分析は、厳密には正しくない。時系列解析では、周期性に関する推測は、古くはピリオドグラムに基づく方法、

Hannan, E.J. Testing for a Jump in the Spectral Function, *J.Royal Statistical Society Ser.B*, **23**, pp.394-404, 1961

あるいは、近年の単位根計量経済学の流れの中で、

Gregoir, S. Efficient Test for the Presence of a Pair of Complex Conjugate Unit Roots in a Real Time Series, *J.Econometrics*, **130**, pp.45-100, 2006

などがある。周期性の分析は、現在も研究が続いているテーマである。

しかしながら、Jevons が周期性を確信したのは、他の 6 系列についても同じようなパターンが見られたためである（図 4）。この点を考察してみよう。

図 4 サイクル内年平均

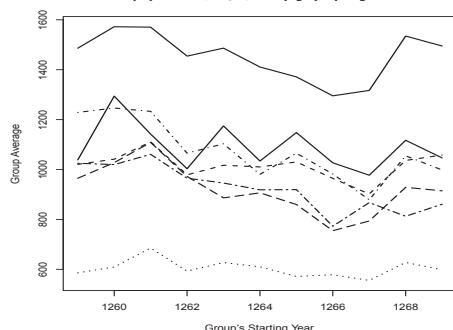


表 5 分散分析表（2 要因モデル）

要因	変動	自由度	不偏分散	F	p -値
穀物	427.4	6	71.23	231.4	0
周期性	35.9	10	3.59	11.67	$8 \cdot 10^{-8}$
誤差	18.5	60	3.08		
全体	481.8	76			

表 3 は、一見すると、先に説明に用いた表 1 に似ているが、表 1 においては最下段に配置されている「サイクル年平均」が、各行に載せられていることに注意する。表 3 のように配置されたデータで、表の縦横ともに何らかの要因に関連しているとみなす場合、これを 2 元配置データと呼ぶ。2 元配置データ (x_{ij}) のモデルは、一般に

$$x_{ij} = \mu + \alpha_i + \beta_j + e_{ij}, \quad i = 1, \dots, p, \quad j = 1, \dots, q \quad (2)$$

と表現される。ここで、 α_i は、行に配置された要因（われわれのケースでは穀物の種類）の「効果」で $\sum_{i=1}^p \alpha_i = 0$ を満たし、 β_j は、列に配置された要因（サイクル内年すなわち周期変動）の「効果」で $\sum_{j=1}^q \beta_j = 0$ を満たすとする。これを 2 要因加法モデル（あるいは単に 2 要因モデル）という。詳細は省略するが、上に述べた 1 要因モデルと同様な分散分析ができる。表 12 に与えられる（穀物 × サイクル内年）2 元表（Jevons の表（表 3）を計算し直したもの。詳細は 8 節）に、このようなモデルを仮定して行った分散分析の結果が、表 5 である。「穀物価格」要因について F の値は極めて大きく、当然ながら、穀物の種類間の平均に差があることを意味している。一方、「周期変動」についても F がかなり大きく、これも有意であることがわかる。つまり表 3 に基づく Jevons の「直観」は、これを 2 要因モデルとみなせば正しかったことになる。

しかし、この結果は「見かけの周期変動」に共通性があることを示しているに過ぎない。真の時系列的な周期性を結論づけるためには、(2)式の β_j で表される「サイクル内変動」が、すべてのサイクルに共通して繰り返されることを調べなくてはならない。

5. 2 要因モデルの解析—繰り返しのあるデータ

2 要因モデルのサイクル内変動の時間的安定性を見るためには、平均に集約された表（表 3）だけではなく、すべてのデータを利用する必要がある。データを一般に x_{ijk} と表す。本例では、 i は穀物の番号、 j はサイクル内の年の番号、 k はサイクルの番号（各要因の組み合わせに対し

て得られる観測値の繰り返しとみなす)を表す。このデータに対して、2要因モデルは

$$x_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}, \quad i = 1, \dots, p, \quad j = 1, \dots, q, \quad k = 1, \dots, r \quad (3)$$

と表される。このモデルに対する分散分析の結果が表6である⁽¹⁷⁾。なお、繰り返しの数を揃えるため、データは欠測のない期間(1276–1377)のみを用い、さらにJevonsの解析と同じく1314年を除いた上で1276から始まる $11 \times 9 = 99$ 年間(1375年まで)のものを用いた。データの総数は $7 \times 99 = 693$ である。

表6 分散分析表(繰り返しのある2要因モデル)

要因	変動 $\times 1/10^6$	自由度	不偏分散 $\times 1/10^3$	F	p -値
穀物(G)	42.45	6	7075.4	41.13	0
周期性(C)	4.65	10	465.4	2.71	0.0060
交互作用(G $\times$ C)	1.59	60	26.6	0.14	1
誤差(E)	114.68	616	186.2		
全体	163.38	692			

周期性要因について、 p -値は0.6%で、表5の結果と比べればかなり大きくなったが、ここでも再び、変動の周期性は1%水準でも有意であるとの結論が出た。個別には、有意にならない(3節の結果)ほどの小さな周期性であるが、(分散分析においては、系列の独立性を仮定しているので)7系列に共通にパターンが見られることは偶然ではあり得ないという結論が得られる。Jevonsが表3によっておそらく意図したであろうこと(すなわち、多くの系列にわたる「平均」操作によって、変動を見いだそうとしたこと)は、この意味では正しかったといえよう。

6. 変数間の相関

標準的な分散分析では、モデルの誤差項 e_{ijk} はすべて独立(かつ同一正規分布)であると仮定する。これに対して、本ケースでは、穀物間の間にかかなり強い相関(これは系列相関ではない)が見られる(表7、表8)。

(17) 表中の、交互作用(interaction)は、「効果の加法性」からの乖離を意味する。ここでは交互作用に対する F の値が非常に小さく、加法モデルがよく当てはまっていると解釈できる。

表 7 穀物価格間の相関

	Wheat	Barley	Oats	Beans	Peas	Vetches	Rye
Wheat	1.000	0.913	0.896	0.877	0.868	0.819	0.924
Barley	0.913	1.000	0.886	0.877	0.881	0.887	0.870
Oats	0.896	0.886	1.000	0.877	0.861	0.835	0.829
Beans	0.877	0.877	0.877	1.000	0.935	0.902	0.870
Peas	0.868	0.881	0.861	0.935	1.000	0.906	0.887
Vetches	0.819	0.887	0.835	0.902	0.906	1.000	0.826
Rye	0.924	0.870	0.829	0.870	0.887	0.826	1.000

このように、変数間に正の相関のある場合、各時点で穀物価格は同じように変動する。変数間の正の相関は、各変数が互いに独立である場合に比べ、 β_j の偶然変動を大きくするように働く。したがって、 β_j 間の差が有意になり易くなる。このような、相関構造を分析に取り込むことは、多変量解析の範疇に入る⁽¹⁸⁾。

ここで、モデルは、(3)と同じ2要因加法モデルとし、誤差項 e_{ijk} について、各穀物変数の間に相関があると仮定する。 $\text{Cov}(e_{ijk}, e_{i'jk}) = \sigma_{ii'}$ とおく。便宜的に(3)式を次のようにベクトルを用いて

$$\mathbf{x}_{jk} = \boldsymbol{\mu} + \beta_j + \mathbf{e}_{jk}, \quad j = 1, \dots, q, \quad k = 1, \dots, r \quad (4)$$

と表す。本例では、添え字 j は「サイクル内の年」の番号を、添え字 k はサイクルの番号を、それぞれ表す。各ベクトルを成分を用いて示すと、

$$\mathbf{x}_{jk} = \begin{pmatrix} x_{1jk} \\ x_{2jk} \\ \vdots \\ x_{pjk} \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} = \begin{pmatrix} \mu + \alpha_1 \\ \mu + \alpha_2 \\ \vdots \\ \mu + \alpha_p \end{pmatrix}, \quad \mathbf{e}_{jk} = \begin{pmatrix} e_{1jk} \\ e_{2jk} \\ \vdots \\ e_{pjk} \end{pmatrix}$$

である。また、誤差項は p -変量正規分布、 $N_p(0, \Sigma)$ にしたがう。 $\Sigma = (\sigma_{ii'})$ は e_{jk} の分散共分散行列である。 Σ がわかっている（既知の）とき、 μ および「周期性の効果」 $\beta_j, j=1, \dots, q$ は、それぞれ

$$\hat{\mu}_i = \frac{1}{T} \sum_{j=1}^q \sum_{k=1}^r x_{ijk}, \quad i = 1, \dots, p, \quad \hat{\beta}_j = \frac{\sum_{i,i'=1}^p \sum_{k=1}^r x_{ijk} \sigma^{ii'}}{r \sum_{i,i'=1}^p \sigma^{ii'}}, \quad j = 1, \dots, q \quad (5)$$

によって、推定される。ここで、 $T=qr$ はデータ長、 $\sigma^{ii'}$ は逆行列 Σ^{-1} の (i, i') 成分である。また、仮説「周期性がない」すなわち $H_0: \beta_1 = \dots = \beta_q = 0$ のための対数尤度比統計量は、

(18) 多変量分散分析。この種の分析は1930年代以降、精力的に研究された。それらの多くは、有名なテキスト Anderson, T.W. *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons Inc., NY, 1958にまとめられている。

$$\lambda = \sum_{j=1}^q \sum_{k=1}^r \sum_{i,i'=1}^p \left\{ (x_{ijk} - \hat{\mu}_i)(x_{i'jk} - \hat{\mu}_{i'}) - (x_{ijk} - \hat{\mu}_i - \hat{\beta}_j)(x_{i'jk} - \hat{\mu}_{i'} - \hat{\beta}_j) \right\} \sigma^{ii'} \quad (6)$$

である。検定は、仮説の下で、統計量 λ が、自由度 $q-1$ の χ^2 分布にしたがうことを利用して行う。

表 8 誤差分散共分散および相関係数の推定値

	Wheat	Barley	Oats	Beans	Peas	Vetches	Rye
Wheat	2.832	0.916	0.896	0.878	0.867	0.824	0.927
Barley	1.780	1.334	0.890	0.884	0.887	0.900	0.881
Oats	0.784	0.534	0.270	0.887	0.863	0.840	0.835
Beans	2.074	1.434	0.647	1.970	0.940	0.909	0.875
Peas	1.849	1.300	0.569	1.674	1.608	0.911	0.899
Vetches	1.789	1.342	0.563	1.646	1.490	1.665	0.844
Rye	2.219	1.448	0.617	1.746	1.620	1.549	2.022

さて、実際には、 Σ は、未知であるので、これをその推定値 $\hat{\Sigma}$ で置き換え、(5)式および(6)を用いる。分散共分散の推定値 $\hat{\Sigma}$ を表 8 に示した。表の対角線以下の部分に分散および共分散(1/100000倍)を、右上の部分に相関係数を、それぞれ示した。

これを用いて得られた β すなわち「サイクル内年効果」の推定値を表 9 に示す。また、各穀物価格の平均 μ は表 10 のとおりである。これらを用いて得られる対数尤度比は $\lambda = 8.03$ であった。自由度は $q-1=10$ であるから $\lambda = 8.03$ に対する p -値 62.6% が得られる⁽¹⁹⁾。よって、サイクル内における平均の差は、全く偶然変動によるものであるといて差し支えない。Jevons はやはり「見かけの周期性」を発見してしまったのである。

表 9 サイクル内年効果

	I	II	III	IV	V	VI	VII	VII	IX	X	XII
$\hat{\beta}$	-13.25	21.65	22.64	0.35	-3.91	-58.86	-19.88	34.78 _I	-15.02	23.61	7.8 _I

表 10 平均穀物価格

	Wheat	Barley	Oats	Beans	Peas	Vetches	Rye
$\hat{\mu}$	1529	1068	633	1104	962	959	1127

(19) ここでは、 Σ が既知である場合の対数尤度比統計量において、 Σ の値を $\hat{\Sigma}$ で代用した。このことは、実際には λ の分布に影響する。多変量解析では、このような問題では、 F の一般化である Wilks の統計量という検定統計量を用いて検定するものとされている。しかしながら、ここでは誤差の自由度は $10 \times 9 = 90$ であるので、推定誤差はあまり大きくない、したがって λ の分布は χ^2 分布で十分よく近似されらる。また、 Σ を推定したことの効果は p -値を大きくする方向に働くから、正しい p -値は 62.6% より (少し) 大きくなるはずである。

7. Jevons の解析 II

論文VIの後半で、Jevons は穀物価格が、サイクルの前半に当たる年で大きくなる傾向があることを指摘し、この事実は（11年を周期とする）規則性すなわち周期性を示唆すると主張した。

原論文の表（論文VI, p.182）を引用する（表11）。表は、それぞれの穀物ごとに、各サイクルでサイクル内の最大（最小）が起こった年のサイクル内年番号を、数え上げたものである。Jevons は表の計数を列ごとに合計した。合計82の計数のうち、第4、5、6、7、8、9の各列（それぞれサイクル内の4年目、5年目...に対応する）の合計が31で、残りの列の合計が51である。後者は、もし周期性がなく、計数が等しい可能性で配分されたときに期待される $\frac{5}{11} \times 82 = 37$ に比べて40% 近くも大きい。Jevons は「これは偶然（accidental）ではない」とだけ述べ、これに対する確率論的評価は加えていない。周期性がなければ、各サイクル内で価格の最大値は各年に均等に分布するから、仮説は $H_0: p=5/11$ である。実際、 p -値は $P(X \geq 51) = 0.0008$ であり、仮説は棄却され、「周期性」が認められることになる。

表11 サイクル内最大値（最小値）を達成するサイクル内年番号の分布

Occurrence of Maxima Prices.

Years.	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th
Wheat -	1	2	2	2	0	1	1	0	1	1	2
Barley -	1	1	1	1	0	2	2	0	1	2	2
Oats -	1	2	1	0	1	1	2	0	0	3	1
Beans -	2	0	3	1	2	1	1	0	0	0	1
Peas -	0	1	5	0	2	1	1	0	0	1	1
Vetches -	1	2	3	1	0	1	0	0	0	1	1
Rye -	0	2	1	1	2	1	0	1	0	1	2
Sums -	6	10	16	6	7	8	7	1	2	9	10

Occurrence of Minima Prices.

Years.	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th
Wheat -	4	1	1	1	1	0	3	0	2	0	0
Barley -	3	2	0	0	1	0	1	1	3	0	2
Oats -	3	1	0	0	0	0	3	0	2	2	1
Beans -	0	2	0	1	0	2	2	0	2	2	0
Peas -	0	2	0	0	2	1	2	2	2	1	0
Vetches -	2	2	0	1	0	1	1	1	0	1	1
Rye -	4	1	0	1	0	2	0	3	0	0	1
Sums -	16	11	1	4	4	6	12	7	11	6	5

現代の視点では、もう少し進んで、表の再下段の計数分布を用いて、一様性の適合度検定を行うことになる。このとき、 $\chi^2=22.1$ で p -値は $P(\chi^2 \geq 22.1)=0.015$ すなわち、有意水準 5% で棄却される（最小値の表に基づいて同じ検定を行うと、 $\chi^2=25.8$, p -値は 0.004 である）。この結果は、2 要因モデルで繰り返しのあるデータに基づく結果（表 5）の p -値に近い。この「サイクル内の最大値の生起時点」に基づく接近法は、時系列の周期性の検証の「ノンパラ」メトリックな方法としてそれほど不合理なものではない。この方法を考案した Jevons の独創性は高く評価されるべきである。

8. 再計算余話

本論文で行ったデータの再解析のために、データを原資料から読み取って入力し計算を行った。データ期間 1259—1400 年のうち、小麦の他のいくつかの穀物価格には欠測がみられ、7 系列すべてが利用できる期間は、1276—1377 年であった。原データを用い、原論文の表（表 3）について、その記述に忠実に（しかしながら原論文には欠測についての言及がないので、これを除いて）計算し直した（表 12）。

表 3 と表 12 を見比べると、原論文の表は、正しい（再計算結果）表を、循環的に右方向に 1 列だけシフトしたものになっていることがわかる。また、対応する数値に微妙な食い違いが見られる。Rogers の原資料は、小数点以下 2 桁の精度で記されているから、この食い違いは丸め誤差の程度を遙かに超えたものである。計算の食い違いはともあれ、表の位置のズレは、信じられないほど大きなうっかりミスであるといえよう。

表 12 「Rogers データ」の周期性（再計算）

	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269
Wheat	1486	1572	1571	1455	1486	1411	1371	1295	1317	1535	1495
Barley	1021	1041	1111	979	1017	1010	1030	964	902	1036	1058
Oats	586	609	685	593	627	610	572	578	556	627	598
Beans	1229	1246	1233	1066	1103	981	1065	982	879	1055	998
Peas	965	1028	1109	972	886	907	860	755	794	929	915
Vetches	1025	1020	1061	965	946	919	920	773	867	813	861
Rye	1039	1294	1142	1003	1175	1034	1148	1027	977	1117	1046
	7350	7811	7911	7032	7241	6871	6966	6375	6291	7112	6972

9. まとめ

本稿では、Jevons が太陽黒点説の根拠のひとつとした「13-14世紀イギリスの穀物価格に見られる11年を周期とした変動」が偶然変動に伴う見かけの変動であることを、現代の分散分析法を用いて明らかにした。

Jevons のデータ解析は、主として2つの点で不適切であった。第1に、Jevons は、1259年から1400年までの年を、11年ごとのサイクルに分け、同一のサイクル内年ごとに平均を求め、それらがサイクル内を変動するパターンを、価格の周期的な変動のパターンだと考えた。これは偶然変動の存在を前提にすると、論理的におかしい。もし価格変動に周期性があれば、長期にわたる繰り返しを平均することで、その周期変動は大数法則によって明瞭になる。逆は、しかしながら正しくない。実際には周期性がなくとも、偶然変動だけからなる見かけの周期性は必ず現れる。この有意性の評価には、同じパターンが繰り返し現れているかどうかを調べる必要がある。「時系列的な平均化」によってその情報が完全に失われしまうので、Jevons の作成した表3では有意性の判断はできない。Jevons は、同じパターンがすべてのサイクルに共通して現れているか否かを検証すべきであった。

ただし、7節で紹介した表は、最大値の発生パターンをサイクルを単位として繰り返し観察したものとみなすことができる。こちらの分析では、上に述べた問題は解決されている。この方法は、一種の「ノンパラ"メトリック」な方法として合理的なものであり、統計学史的にもう少し注目されて良いと思われる。

第2は、このパターンが7種の穀物に共通に見られることに関連する。Jevons は、おそらく、この事実は（偶然変動であればパターンが共通に現れる可能性は少ないから）そのようなパターンは周期性に関連するものであると考えた、と思われる。各穀物価格が相互に独立に変動しているならば、この推論は正しい。しかしながら、「Rogers データ」の各変数間には強い相関があり、（互いに似た（偶然）変動をするから）各変数の「見かけのパターン」が類似のものである確率は高くなる。

変数間の従属性は Jevons の第2の方法においても問題となるが、Jevons は、これについては、問題を正しく認識していることがわかる（論文VI, p.183）。

The results which I have given would be altogether conclusive as to the influence of the sun-spot if each of the seven commodities varied independently of the rest. As they are, however, not independent I must regard these results as of a merely provisional value, showing the need of further investigations of the same nature, which will be surrounded by difficulties.

この「従属性」の問題は、しかしながら、Jevons が想定したよりも遙かに深刻だったのである。本稿では、2要因モデルの分散分析という視点で、Jevons の解析を再検討した。ここでは、

誤差項の時系列的な相関は考慮していない。時系列解析的な接近法をとれば、図4に見られる（一瞥すると）正弦曲線的なパターン—これが Jevons に周期変動を錯覚させたもう一つの原因であると想像されるが—が現れる理由が解明されるかもしれない。これは今後の課題としたい。

10. 補足（1 要因モデルの分散分析）

データの配列を x_{ij} , $i=1, \dots, r, j=1, \dots, p$ とする。このデータのモデルを

$$x_{ij} = \mu + \alpha_j + e_{ij}, \quad i = 1, \dots, r, j = 1, \dots, p \quad (7)$$

とする。誤差 e_{ij} はすべて互いに独立に平均 0, 分散 σ^2 の正規分布にしたがう確率変数であると仮定する。 μ は全体平均, $\alpha_j, j=1, \dots, p$ は「要因効果」を表す母数である。このようなデータは、1 元配置データと呼ばれ、(7)は 1 要因モデルと呼ばれる。

全体の平均を $\bar{x}$, 列平均を $\bar{x}_{.j}, j=1, \dots, p$ とする。このとき、全変動 $T = \sum_{i,j} (x_{ij} - \bar{x})^2$ は

$$T = r \sum_{j=1}^p (\bar{x}_{.j} - \bar{x})^2 + \sum_{j=1}^p \sum_{i=1}^r (x_{ij} - \bar{x}_{.j})^2$$

と分解される。上の右辺第 1 項は「グループ間変動」、第 2 項は「グループ内変動」であり、分散分析表の「変動」の列に記される。第 1 項をその自由度 $p-1$ で割って得られる V_A は、仮説（グループ平均に差がない）の下では、誤差分散 $\sigma^2 = Ee_{ij}^2$ の不偏な推定値であり、グループ間に差があれば大きくなる。したがって、 V_A の大小によって仮説の妥当性が判断できる。第 2 項をその自由度 $p \times (r-1)$ で割って得られる V_E は、グループ平均の差の有無にかかわらず誤差項 e_{ij} の分散の不偏推定値を与える（これらは「不偏分散」の列に記される）。このような理由で、検定統計量は $F_A = V_A/V_E$ とされる。 F_A の分布は、 $(p-1, (p-1)r)$ を自由度とする F 分布となる。この分布に基づいて得られた F_A に対する p -値を求め、グループ間変動の大きさ（すなわちグループ間の平均の差の有意性）を判断する。