RSS Higher Certificate in Statistics, 2010

Module 7 : Time series and index numbers

Solutions

Question 1

(i) 加法分解は$Y=C+S+I$となる．季節調整前のデータに変換を適用するべきでない．
乗法分解は$Y=C\times S\times I$となる．季節調整前のデータに対数変換を適用するべきだ．

以下のように分解を選択すればよい. 系列の予備知識が使えるかもしれない. 加法分解は，系列自体が減法で得られる場合や 0 または負の値を含む場合に適していて，乗法分解は，インデックスや経済的な系列の場合に適しているといえる．われわれは関連する系列（調査中の系列の成分や集計値，以前の系列の値など）を知っているかもしれない．また，各分解を用いて季節調整モデルにあてはめ，どちらのあてはめがよりよいかを（目で視覚的に，または診断を見直すことによって定量的に）見ることもできる．

季節調整系列は$Y=C+I$または$Y=C\times I$となる．（最初にCを取り除くことによって）季節成分Sを推定し，その後（加法分解ならば引き算，乗法分解ならば割り算で）YからそのSを除去することでこの式は得られる．

(ii) 系列へのイースターの効果があれば，すなわち，特にイースターのために系列が大きくなったり小さくなったりする場合は，3 月と 4 月の間のイースターの動向は問題を引き起こす可能性がある．これはよくあるケースだ．たとえば，通常の取引日が少ないためにいくらかの系列が小さくなる一方で，イースターで強化される何らかの形の取引に関わっている他の系列は大きくなる．このようなケースでは系列の性質の変化によってではなく，単にカレンダーによって動向がゆがめられる．たとえば，ある年はイースターが 3 月にあり他のある年は 4 月にあるという理由だけで，前者の 2 月と 3 月の間の伸びは後者の伸びよりも小さくなりうる．このことは，系列の増加/減少に対し誤解を招き，系列の動向に基づいて決定をする人に対して問題を引き起こすかもしれない．

イースターの効果は，イースターの調整がある場合とない場合の季節調整をし，相対的な調整を目で見たり診断を検討したりして評価することによって検定される．その診断には，（不規則成分に対して回帰したときに）イースター自体に対するダミー変数のパラメータ推定の有意性が含まれているだろう．

イースターの効果があるならば，それを推定して季節調整前に系列から除去する――たとえば，イースターがもたらす差の半分を“イースターの月”から除去し“イースターのない月”に加える――ことによって調整が行われる. イースターの調整が永久的であること，つまり季節調整後にこの調整は残らないことに注意せよ．

カレンダーに関連した他の影響として，取引日や閏年，他にも旧正月，5 月の銀行休日，ラマダーンなどの移動する休日が含まれる．クリスマスが何曜日になるかも系列に影響するかもしれない．

(iii) 2004 年 10 月のものと 2005 年 8 月のものの，2 つの加法の外れ値が確認されている．これらは通常の季節変動では説明されていない効果であり，永久的な変化を示すようには見られるわけではなく 1 回限りの繰り返さないイベントである．外れ値は 2 つとも 5%有意水準で有意である——これは，この系列の季節調整に影響を与えているという真の証拠があることを意味している．イースターを検定した，すなわち，3 月と 4 月の間のイースターの動きが系列に影響を与えているかどうかを検定したが，パラメータは 0 との差が有意ではないのでイースターはこの系列に影響を与えていないことがわかる．

両外れ値は，このような現象を扱う手段に応じて調整される．すべての外れ値が調整される方法かもしれないし，実世界のある種のイベントに一致するように見える外れ値のみが調整される方法かもしれない．イースターに対しては調整すべきでない．

両者の違いは，（(ii)の解答中で述べたように）イースターの調整は永久的である——季節調整が行われた後に残らない一方で，外れ値の調整は一時的である——季節調整が行われた後に残るということだ．

Question 2

(i) 未調整の時系列の移動平均は，データを平滑化し，トレンド成分の初期推定量を得るのに用いられる．トレンドのこの推定量はその後，季節成分と不規則成分の組み合わせを残すために未調整の時系列データから除去することができる．結果として得られる時系列は，季節成分の推定量を得る他の移動平均によって，1 度にひと月または四半期を平滑化する．この季節変動は未調整の時系列から順に除去することができ，改良されたトレンド成分の推定量を得るため平滑化された季節調整済時系列の推定量が残る．

(ii) 2005 年の Q3 に対する対称な 5 点単純移動平均は

$$\frac{50.0+36.5+43.0+44.5+38.9}{5}=42.58$$

で与えられ，他も同様に求められる．この計算は最初の 2 つの四半期と最後の 2 つの四半期に対して行うことはできない．以下のようになる．

商品の生産量 *y*

Period	2005 Q1	2005 Q2	2005 Q3	2005 Q4	2006 Q1	2006 Q2	2006 Q3	2006 Q4	2007 Q1	2007 Q2	2007 Q3	2007 Q4
Production	50.0	36.5	43.0	44.5	38.9	38.1	32.6	38.7	41.7	41.1	40.5	33.8
Moving av.			42.58	40.20	39.42	38.56	38.00	38.44	38.92	39.16		

対称な（均等または不均等な加重）移動平均は，注目している点の前後に等距離離れた系列の値に同程度の重要性を与えるが，このことはたいてい直感的に理に適っている．だが，特にこの系列について上述したように，系列の最初や最後においてはこのような移動平均を計算できない．たとえば，この系列には 12 のデータ項目があるものの，平滑化した系列には 8 項目しかない．対称な移動平均は単に，系列の両端で使用できない．

系列の両端の領域でも平滑化された推定量を得るには，非対称の移動平均を使用すればよい．またどのようなデータが利用できるかによって，点の片側からのデータをもう一方のデータよりも多く用いて決められた平均を使用して，“中心から外れて” 平滑化された値を計算する．または，モデリング技法を長期系列で使用した後，移動平均を用いて系列を補外法で推定してもよい．

(iii) 定義より，均等な加重移動平均は，計算に含まれる全ての点に等しい重みを与える．これは多くの場合，注目している点から“近い”点に対して遠く離れた点に重みを与えすぎていることを意味する．“近い”点は注目している点からはより“影響を受け”やすく，それに比べ，系列の自然変動からは影響を受けにくい．したがって，“近い”点により大きな重みを与える不均等な重みはしばしば好ましい．不均等な重みのもう一つの特徴として，トレンドの転換点や変曲点を効率的に取り入れた多項式を近似するように選ぶことができる．

(iv) 長い移動平均は一般的に，より多くのデータポイント数からの情報を取り入れているので，

得られるトレンドはより滑らかになる．

しかしながら平均を長くすることには不利な点がある．線形的な傾向は短期間は続くであろうが長期間にわたって続くことはほとんどない．トレンドが曲がっている場合，使う平均が長すぎると歪んでしまう．このように移動平均の長さの選択は，平滑化とトレンドを追いかけるという 2 つの目的のバランスによるだろう．

Question 3

(i)

$$\frac{P_L(Jan,Mar)}{P_L(Jan,Feb)}=\frac{\dfrac{\sum_i p_{Mar,i}q_{Jan,i}}{\sum_i p_{Jan,i}q_{Jan,i}}}{\dfrac{\sum_i p_{Feb,i}q_{Jan,i}}{\sum_i p_{Jan,i}q_{Jan,i}}}=\frac{\sum_i p_{Mar,i}q_{Jan,i}}{\sum_i p_{Feb,i}q_{Jan,i}}$$

p は価格を，q は数量を表していて，i で各商品を識別している．

(ii)

$$P_{Lo}(Feb,Mar;Jan)=\frac{\sum_i p_{Mar,i}q_{Jan,i}}{\sum_i p_{Feb,i}q_{Jan,i}}$$

$$=\frac{\sum_i p_{Mar,i}\dfrac{p_{Feb,i}}{p_{Feb,i}}\dfrac{p_{Jan,i}}{p_{Jan,i}}q_{Jan,i}}{\sum_i p_{Feb,i}\dfrac{p_{Jan,i}}{p_{Jan,i}}q_{Jan,i}}=\frac{\sum_i R_{Feb,Mar,i}R_{Jan,Feb,i}v_{Jan,i}}{\sum_i R_{Jan,Feb,i}v_{Jan,i}}$$

R は価格の相対比を，v は価格で表した名目的な価値を表している．

ロウ価格指数の重みは $R_{Jan,Feb,i}v_{Jan,i}$ である．これらは，1 月から 2 月までの価格変動によって増えた 1 月の値である．

(iii)

$$\frac{P_P(Jan,Mar)}{P_P(Jan,Feb)}=\frac{\dfrac{\sum_i p_{Mar,i}q_{Mar,i}}{\sum_i p_{Jan,i}q_{Mar,i}}}{\dfrac{\sum_i p_{Feb,i}q_{Feb,i}}{\sum_i p_{Jan,i}q_{Feb,i}}}$$

(iv)

$$\frac{P_P(Jan,Mar)}{P_P(Jan,Feb)}=\frac{\sum_i v_{Mar,i}}{\sum_i v_{Feb,i}}\frac{\sum_i p_{Jan,i}q_{Feb,i}}{\sum_i p_{Jan,i}q_{Mar,i}}=\frac{\dfrac{\sum_i v_{Mar,i}}{\sum_i v_{Feb,i}}}{Q_{Lo}(Feb,Mar;Jan)}$$

よって，値の伸びをロウ数量指数で割っている．

(v)

$$\frac{P_F\left(Jan, Mar\right)}{P_F\left(Jan, Feb\right)} = \frac{\sqrt{P_L\left(Jan, Mar\right) P_P\left(Jan, Mar\right)}}{\sqrt{P_L\left(Jan, Feb\right) P_P\left(Jan, Feb\right)}}$$

$$= \sqrt{\frac{P_L\left(Jan, Mar\right) P_P\left(Jan, Mar\right)}{P_L\left(Jan, Feb\right) P_P\left(Jan, Feb\right)}}$$

$$= \sqrt{\frac{\sum_i p_{Mar,i} q_{Jan,i}}{\sum_i p_{Feb,i} q_{Jan,i}} \times \frac{\frac{\sum_i v_{Mar,i}}{\sum_i v_{Feb,i}}}{Q_{Lo}\left(Feb, Mar; Jan\right)}}$$

$$= \sqrt{\frac{\sum_i v_{Mar,i}}{\sum_i v_{Feb,i}} \frac{P_{Lo}\left(Feb, Mar; Jan\right)}{Q_{Lo}\left(Feb, Mar; Jan\right)}}$$

が(i),(ii)及び(iv)よりわかる.

Question 4

(i) ラスパイレスの出生率指数の式は，r を出生率，w を女性の人数，i を年齢群とし，0 を基準時点，t を現時点としたとき，

$$\frac{\sum_i r_{ti} w_{0i}}{\sum_i r_{0i} w_{0i}}$$

となる．また，b を出生児数とする．

すると，上の式は他にも $\dfrac{\sum_i r_{ti} w_{0i}}{\sum_i b_{0i}}$， $\dfrac{\sum_i \frac{r_{ti}}{r_{0i}} b_{0i}}{\sum_i b_{0i}}$ と書き表せる．

パーシェの出生率指数の式は，同じ表記を用いて

$$\frac{\sum_i r_{ti} w_{ti}}{\sum_i r_{0i} w_{ti}}$$

と表せる．

これも他に $\dfrac{\sum_i b_{ti}}{\sum_i r_{0i} w_{ti}}$， $\dfrac{\sum_i b_{ti}}{\sum_i \frac{r_{0i}}{r_{ti}} b_{ti}}$ と書き表せる．

(ii)

Age group	*Number of women in 2008*	*Number of live births in 2008*	*Number of women in 2009*	*Number of live births in 2009*
15–17	3095	30	2985	28
18–21	4056	439	4027	401
22–29	7483	984	7514	975
30–39	10247	426	10473	453
40–49	9835	37	9626	52

（上記の基本式を用いて計算した）ラスパイレスの出生率指数は

$$\frac{\frac{28}{2985}\times 3905+\frac{401}{4027}\times 4056+\frac{975}{7514}\times 7483+\frac{453}{10473}\times 10247+\frac{52}{9626}\times 9835}{\frac{30}{3095}\times 3905+\frac{439}{4056}\times 4056+\frac{984}{7483}\times 7483+\frac{426}{10247}\times 10247+\frac{37}{9835}\times 9835}$$

$$=\frac{1900.25}{1916}=0.9918 \qquad \text{または}99.18\%.$$

（上記の基本式を用いて計算した）パーシェの出生率指数は

$$\frac{\frac{28}{2985}\times 2985+\frac{401}{4027}\times 4027+\frac{975}{7514}\times 7514+\frac{453}{10473}\times 10473+\frac{52}{9626}\times 9626}{\frac{30}{3095}\times 2985+\frac{439}{4056}\times 4027+\frac{984}{7483}\times 7514+\frac{426}{10247}\times 10473+\frac{37}{9835}\times 9626}$$

$$=\frac{1909}{1924.48}=0.9920 \qquad \text{または}99.20\%.$$

フィッシャーの出生率指数は $\sqrt{0.9918\cdot 0.9920}=0.9919$ または 99.20%.

(iii) 女性の人数の指数は，各年齢群の女性の人数の変化の，対応する出生率で重みづけした加重平均である．これは，女性の人数全体の変化が出生児数の変化にどう貢献してきたか指標を与えようとしている，とみなすことができる．

このように指数を計算すると，

$$\frac{\sum_i w_{ti} r_{0i}}{\sum_i w_{0i} r_{0i}}$$

(w は女性の人数，r は出生率，i は年齢群を示し，0 を基準時点，t を現時点とする)

$$=\frac{2985\times\frac{30}{3095}+4027\times\frac{439}{4056}+7514\times\frac{984}{7483}+10473\times\frac{426}{10247}+9626\times\frac{37}{9835}}{3095\times\frac{30}{3095}+4056\times\frac{439}{4056}+7483\times\frac{984}{7483}+10247\times\frac{426}{10247}+9835\times\frac{37}{9835}}$$

$$=\frac{1924.4807}{1916}=1.0044 \qquad \text{または}100.27\%$$