第９章　結果の推定方法と標本誤差等

労働力調査は標本調査であることから，その結果や誤差については統計理論に基づいて推定される。また，季節性を除去する方法として季節調整等がある。本章では，労働力調査における結果の推定方法と標本誤差のほか，季節調整等について解説する。

１　線型推定

標本調査は，一部を調査して全体を推定しようとするものであるが，全体の推定値は，標本から得られた値に，抽出率（抽出確率－標本として選ばれる確率）の逆数を乗じることによって得ることができる。このような推定を線型推定という。

労働力調査の場合に当てはめれば，調査区の抽出の際に層別抽出を行っているので，まず各層で独立に推定を行い，次に，各層における推定値を足し合わせて全体の推定値を得ることになる。第l層における推定方法について，第l層の就業者数X_lを推定する場合を例に説明することにする。

第l層で抽出された調査区がm_l個，各調査区のウエイト注)が$w_{li}(i = 1, 2, \cdots, m_l)$であるとし，第$i$調査区において，抽出された住戸全体で就業者が$X_{li}(i = 1, 2, \cdots, m_l)$ 人居住していたとすると，線型推定値は次のようにして求めることができる。

① まず，抽出された調査区内の就業者数の合計を推定する。抽出率はウエイトの逆数としていたから，抽出率の逆数とはウエイトそのものになる。したがって，$X_{li}w_{li}$が第i調査区の就業者数になる。これは，ウエイト３の調査区の場合，三つに一つの割合で住戸を調査するから，調査した住戸に居住する就業者の数が50人であったならば，50×３＝150人をその調査区の就業者数と推定するというものである。別の見方をすれば，抽出された住戸は，３住戸分を「代表」しているのであるから，各住戸の就業者数を３倍し，それを足し合わせれば推定値が得られるともいうことができる。この「代表」の度合いがつまり抽出率の逆数なのである。

② 次に，層全体の就業者数X_lの推定値$\hat{X}_l$を求める。調査区の抽出は，確率比例抽出であったから，各調査区の抽出確率は等しくはなく，したがって推定値も同じ値を一律に乗じて求めるというわけにはいかない。そこで，同じ値を乗じて求める代わりに，調査区内の就業者数に調査区の抽出確率の逆数を乗じて推定値を求めることとする。

第i調査区の就業者数は$X_{li}w_{li}$と推定されていることから，第i調査区一つから層内全体の就業者数$\hat{X}_l$を推定しようとするとき，層内の全調査区のウ

注）第８章の２を参照

エイトの合計をw_l注）とすれば，第i調査区からの推定値$\hat{X}_{li}$は，

$$\hat{X}_{li} = (X_{li}w_{li}) \times \frac{w_l}{w_{li}} = X_{li}w_l$$

となり，第i調査区のウエイトw_{li}によらない値となる。このため，$\hat{X}_{li}(i = 1,\ 2,\ \cdots,\ m_l)$の平均値をこの層の推定値と考えることができるから，層全体の就業者数$\hat{X}_l$は，

$$\hat{X}_l = \frac{1}{m_l}\sum_{i=1}^{m_l}\hat{X}_{li} = \frac{1}{m_l}\sum_{i=1}^{m_l}X_{li}\,w_l = \frac{w_l}{m_l}\sum_{i=1}^{m_l}X_{li}$$

と推定できる。ウエイトの逆数を抽出率としたことにより，上の式の$\frac{w_l}{m_l}$は調査区によらない定数になっている。この$\frac{w_l}{m_l}$を線型推定用乗率という。

結局，第l層の就業者の推定式は，各調査区中の就業者数を単純に足し合わせ，$\frac{w_l}{m_l}$倍すればよいという簡単な式になっている。

２　比推定の考え方

１で述べたように，線型推定により推定値は得られるのであるが，補助的な情報を利用することにより，精度を高めることができる。

上の例で考えてみると，第l層の就業者数X_lの線型推定値$\hat{X}_l$を求める方法と全く同様な方法により，第l層の総人口P_lの線型推定値$\hat{P}_l$を求めることができる。すなわちi番目の標本調査区において抽出された住戸にP_{li}人の者が居住していたとすると，

$$\hat{P}_l = \frac{w_l}{m_l}\sum_{i=1}^{m_l}P_{li}$$

となる。$\hat{P}_l$も$\hat{X}_l$も標本からの推定値であるから，標本の選ばれ方によって，実際の値であるP_lやX_lより大きくなったり小さくなったりする。しかし，この二つの推定値の実際の値からのずれ方は，同じ方向であることが多いと考えられる。というのは，例えば世帯規模の大きい世帯からなる調査区がたまたま数多く抽出された場合，$\hat{P}_l$も大きい値となるが，同時に$\hat{X}_l$も大きい値となる可能性が高いからである。その結果，$\hat{P}_l$と$\hat{X}_l$の比を考えると，$\hat{P}_l$や$\hat{X}_l$そのものに比べ，かなり安定することが予想される。

そこで，仮に人口の大きさP_lが，別の資料により正確に知ることができたとすると，$\hat{X}_l$そのものを推定値とするより，

注）国勢調査時に得られる層内の全調査区のウエイトの合計

$$\tilde{X} = P_l \times \frac{\hat{X}_l}{\hat{P}_l} = \hat{X}_l \times \frac{P_l}{\hat{P}_l}$$

を推定値とした方が安定した値を得ることができる。この方法は比推定と呼ばれ，P_lをベンチマーク人口という。この方法を用いると，$\hat{X}_l$の誤差は，$\hat{P}_l$の誤差が同じ方向に向かう場合，かなり縮小する。このように，別途正確な数値が得られるものと高い正の相関を持つものの推定には，比推定は非常に有効である。

線型推定値は標本調査区のウエイトの情報など国勢調査結果に基づく推定値であることから，国勢調査結果以降の人口の移動等は加味していないため，実際の推定値との間に乖離が生じている可能性がある。そのため，労働力調査では線形推定値を求めた後，比推定を用いることでその乖離を補正している。

３　推定方法

(1) ベンチマーク人口の推計方法

ベンチマーク人口（以下「基準人口」という。）としては，総務省統計局が毎月公表している「推計人口」を利用している。推計人口とは，国勢調査による人口を基準とし，次の国勢調査結果が得られるまでの間，その後の出生児数・死亡者数（人口動態統計）及び出国者数・入国者数（出入国管理統計）のデータにより男女・年齢階級別に毎月１日現在の人口を推計したものである。年齢階級別人口は，国勢調査時点を出発点として，次の式により毎月計算される。

$$\begin{aligned}
\left(\begin{array}{l}\text{今月１日現在の}\\ \text{当該年齢階級人口}\end{array}\right) &= \left(\begin{array}{l}\text{前月１日現在の}\\ \text{当該年齢階級人口}\end{array}\right) - \left(\begin{array}{l}\text{前月1日現在の当該}\\ \text{年齢階級人口のうち}\\ \text{前月中に死亡した者}\end{array}\right) \\
&\quad + \left(\begin{array}{l}\text{前月中新たに当該年}\\ \text{齢階級に達した人口}\end{array}\right) - \left(\begin{array}{l}\text{前月中新たに一つ上の}\\ \text{年齢階級に達した人口}\end{array}\right) \\
&\quad + \left(\begin{array}{l}\text{前月中の当該年齢}\\ \text{階級入国者数}\end{array}\right) - \left(\begin{array}{l}\text{前月の当該年齢}\\ \text{階級出国者数}\end{array}\right)
\end{aligned}$$

※　15歳以上人口が推計の対象であるため，上式では出生要因を考える必要はない。

なお，労働力調査では，地域別の集計を行っているため,都道府県別の推計人口も別途必要となる。

(2) 基本集計

ア　全国結果の推定方法

就業者数や完全失業者数等の調査結果の算出は，大都市部・非大都市部，男女，年齢５歳階級（15区分注)）別人口を基準人口とする比推定によっている。算出の基本的考え方は，以下のとおりである。

① 各標本調査区の男女，年齢階級別調査人口に線型推定用乗率を乗じて必要な合算を行い，男女，大都市部・非大都市部，年齢階級別人口の線型推定値を算出する。

② 男女，大都市部・非大都市部，年齢階級別に，基準人口をそれぞれ①で算出した線型推定値で除し，比推定用乗率を算出する。

③ 各標本調査区の属性Xを有する男女，年齢階級別調査人口に，線型推定用乗率を乗じて必要な合算を行い，さらに②で算出した比推定用乗率を乗じて，男女，大都市部・非大都市部，年齢階級別の比推定値$\tilde{X}$を算定する。

④ この比推定値$\tilde{X}$を大都市部と非大都市部について合算した後，必要に応じて，男女，年齢階級について合算して就業者数や完全失業者数等の結果数値を得る。

（参考）上記①，②，③をまとめて計算式で表すと，次のとおりである。

$$\tilde{X} = \sum_{l=1}^{L} \frac{1}{m_l} \sum_{i=1}^{m_l} \frac{w_l}{w_{li}} \cdot f_{li} \cdot x_{li} \frac{P}{\sum_{l=1}^{L} \frac{1}{m_l} \sum_{i=1}^{m_l} \frac{w_l}{w_{li}} \cdot f_{li} \cdot P_{li}}$$

$$= \sum_{l=1}^{L} \sum_{i=1}^{m_l} x_{li} F_l \frac{P}{\sum_{l=1}^{L} \sum_{i=1}^{m_l} P_{li} \cdot F_l}$$

上記計算式のうち，

$l(=1,\ 2,\ \cdots,\ L)$は10地域，層による区分の番号

$i(=1,\ 2,\ \cdots,\ m_l)$は各区分中の標本調査区の番号

X_{li}は第l区分，第i標本調査区内の属性Xを有する（男女，年齢階級別）調査人口

w_{li}は第l区分，第i標本調査区のウエイト

f_{li}は第l区分，第i標本調査区の住戸の抽出率の逆数（$= w_{li}$）

w_lは第l区分に含まれる全ての調査区のウエイトの合計

注）15～19 歳から 80～84 歳までの５歳階級及び 85 歳以上

m_lは第l区分の標本調査区数

F_lは第l区分の線型推定用乗率（$= w_l/m_l$）

Pは（男女，大都市部・非大都市部，年齢階級別）基準人口

P_{li}は第l区分，第i標本調査区内の（男女，年齢階級別）調査人口

$\frac{P}{\sum_{l=1}^{L}\sum_{i=1}^{m_l} P_{li}\cdot F_l}$は比推定用乗率を表す。

※　なお，実際の集計に当たっては，あらかじめ線型推定用乗率に比推定用乗率を乗じた集計用乗率を作成して各個票データに付与し，これを合算することによって結果数値を得ている。

イ　地域別結果の推定方法

10地域注1)別結果は，上記の全国結果と同様の方法で算出した後，10地域の合計が全国結果と一致するように補正している。

(3) 詳細集計

ア　全国結果の推定方法

全国結果の算出は基本集計と同様の方法で行っているが，比推定に用いる基準人口注2)については，基本集計の調査結果に基づく男女，年齢10歳階級（５区分注3)），就業状態（就業者，完全失業者，非労働力人口）別人口としている。

イ　地域別結果の推定方法

まず，月次結果について，10地域，男女，年齢５歳階級（13区分注4)）別に，国勢調査に基づく推計人口を基準人口とする比推定により算出する。

次に該当する一年間の月次結果を単純平均して年平均値を算出し，さらに，基本集計の年平均結果の10地域，男女，年齢10歳階級（５区分注3)），就業状態（就業者，完全失業者，非労働力人口）別人口を基準人口とする比推定によって地域別年平均結果を算出する。

なお，計算途中の線型推定値は，基本集計結果の算出の際に用いた線型推定用乗率による集計値である。

注１）第８章の２を参照

注２）詳細集計では，雇用・失業情勢の詳細な状況を調査することを目的としていることから，自衛隊営舎内（艦船内）居住者及び刑務所等の矯正施設収容者は集計の対象から除外している。

注３）15～24 歳から 45～54 歳までの 10 歳階級及び 55 歳以上

注４）15～19 歳から 70～74 歳までの５歳階級及び 75 歳以上

４　推定値の誤差

例えば，ある時点における就業者数を，統計調査によって推定しようとする場合，結果数値が必ずしも真の値に一致するわけではない。この差を誤差という。結果数値をみる場合，誤差の存在を常に認識しておく必要がある。

また，通常，誤差は，標本調査であることに起因する「標本誤差」とそれ以外の実地調査における調査票の誤記入などに起因する「非標本誤差」に分けて考えられている。

(1)　標本誤差

労働力調査の場合，国内居住者全員を調べているわけではなく，その一部である標本を調べて全体を推定している。第８章及び本章の３まででみたように，労働力調査の標本抽出方法及び結果の推定方法はやや複雑であるが，単純化して考えれば，抽出率が1/1000で10万人調査した場合，そのうち５万人が就業者であったとすると，

$$5\text{万人} \times \frac{1000}{1} = 5000\text{万人}$$

という算式によって就業者数は5000万人であると推定しているといってよい。

このとき，この5000万人という結果数値は真の値に等しいとは限らない。なぜなら，調査された10万人という集団は全国民の完全な縮図とは限らないからである。同じ時点において標本の抽出をやり直して10万人調査できたとしても，再び10万人中５万人が就業者数となるとは限らず，５万1000人かもしれないし，４万9000人かもしれない。つまり，真の値は一つであっても，推定値はそれより大きくなっているかもしれないし，小さくなっているかもしれないのである。このように，標本から推定することによって生ずる誤差を標本誤差という。

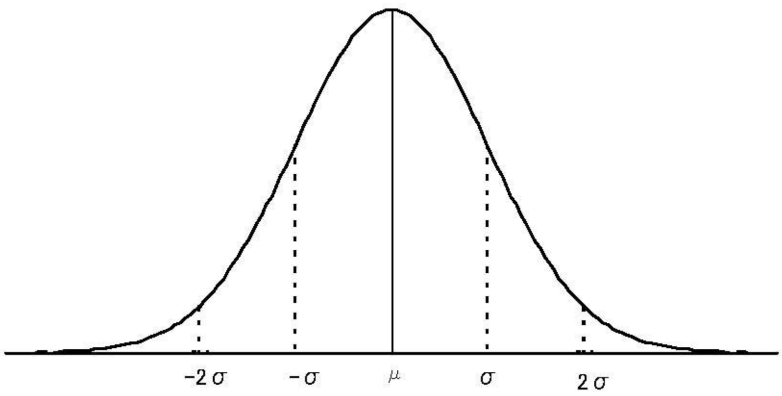
労働力調査の結果は，このような標本誤差が避け得ないものであるから，結果数値をみる場合には注意が必要である。例えば，先月5000万人であった就業者数が今月5001万人になったとした場合，仮に後に述べるような季節性がないとしても，これをもって直ちに就業者数が増加したとは判断できない。先月の数値も今月の数値も共に標本誤差を含んでいるから，真の値は逆に減少しているかもしれないからである。

しかし，標本誤差の存在は，必ずしも結果数値が信頼性を持たないということにはつながらない。というのは，推定値が真の値に「近い」ということは事実であり，また大きく真の値から離れることは少ないということもいえるからである。したがって，上の例でいえば，もし5000万人から5100万人に増加していれば，真の値も増加しているであろうということはかなりの確率でいえるのである。このような判断に根拠を与えるものとして標本理論がある。

標本理論は，標本調査から得られた推定値が真の値からどの程度離れる可能性があるかを理論的に示してくれる。このとき，真の値からの距離を測定する際の物差しになるのが「標準誤差」である。抽出を何度も繰り返し推定を行えば，推定値は真の値の周りにある分布を示すであろう。この分布の広がりが小さければ精度の良い推定ということができる。分布の広がり具合は，通常，分布の標準偏差で示され，この標準偏差を標本理論では標準誤差と呼んでいる。標準誤差は精度を示す指標であると同時に，誤差を測る尺度となる。また，標準誤差を真の値に対する比率で示したものを標準誤差率という。

標本理論によれば，推定値$\bar{x}$は真の値μの周りにほぼ正規分布をしていると考えられ（図９－１），標準誤差σが分かっているとすると，$\bar{x}$とμの差がσ未満となる確率すなわち$|\bar{x}-\mu|<\sigma$となる確率は約68%，$|\bar{x}-\mu|<2\sigma$となる確率は約95%であることがいえる。つまり，推定値の誤差は，３回中２回はσの範囲に収まっており，2σの範囲を超えることは20回に１回程度しかないことがいえるのである。

図９－１ 推定値の分布



労働力調査の場合，ある月の推定値が5000万人のとき，標準誤差は約25万人であると推定されているから，調査結果を正確に記述しようとするなら，単に推定値が5000万人であるというのではなく，例えば5000万人±25万人の間に真の値が2/3の確率で存在するというように記述する必要があろう。

標準誤差は，標本の抽出方法あるいは結果の推定方法が複雑な場合，簡単には求められないが，10万人を無作為に抽出したというように考えると，労働力調査のような人数の推定の場合，15歳以上人口をN，ある属性を持つ人口をX，標本数nの標本によるXの推定値を$\bar{x}$としたとき，$\bar{x}$の標準誤差$\sigma(\bar{x})$は，

$$\sigma(\bar{x}) \fallingdotseq N \times \sqrt{\frac{p(1-p)}{n}} \qquad p = \frac{X}{N}$$（Xの15歳以上人口に占める割合）

となり，標準誤差$\sigma(\bar{x})$を真の値Xで割った標準誤差率は，

$$\frac{\sigma(\bar{x})}{X} \fallingdotseq \frac{N\sqrt{\dfrac{p(1-p)}{n}}}{X} = \frac{N\sqrt{\dfrac{p(1-p)}{n}}}{Np} = \sqrt{\frac{1-p}{pn}}$$

となる。この式を見て分かるように，標準誤差は標本数の平方根に反比例して小さくなり，標本数が４倍になれば標準誤差は半分になる。また，全体に占める割合pが小さい場合，標準誤差は小さくなるが，標準誤差率は逆に大きくなる。

労働力調査の実際の標準誤差は，実務上の制約などから２段抽出法を採っているため，上の式で示される値よりはやや大きくなっている。

ア　全国結果の推定値の大きさ別標準誤差

月別値及び年平均値の標準誤差率は，標本の交代を行うために設けられた８組の副標本を利用し，下記の算式により推定されたものである（ただし，組別の推定値が独立に正規分布していると仮定している。）。

年平均値用

$$\sqrt{\frac{1}{8(8-1)}\sum_{k=1}^{8}\left(\bar{\bar{X}}_k-\bar{\bar{X}}\right)^2} \Big/ \bar{\bar{X}}$$

ここで，$\bar{\bar{X}}_k$は第k副標本による，$\bar{\bar{X}}$は全標本による，属性Xを有する人口の推定値の年平均値を表す。

月別値用

$$\sqrt{\frac{1}{8(8-1)}\sum_{k=1}^{8}\left(\tilde{X}_k-\tilde{X}\right)^2} \Big/ \tilde{X}$$

ここで$\tilde{X}_k$は第k副標本による，$\tilde{X}$は全標本による，属性Xを有する人口の月別推定値を表す。

属性ごとの標準誤差率を曲線の当てはめにより平均的に評価し，推定値の大きさ別におおよその標準誤差及び標準誤差率を示すと次のようになっている。

(ｱ) 基本集計

年平均の結果の標準誤差

推定値の大きさ（万人）	標準誤差（万人）	標準誤差率（%）
5000	11.0	0.2
2000	7.0	0.3
1000	5.0	0.5
500	3.5	0.7
200	2.2	1.1
100	1.6	1.6
50	1.1	2.3
20	0.7	3.6
10	0.5	5.1

毎月の結果の標準誤差

推定値の大きさ（万人）	標準誤差（万人）	標準誤差率（%）
5000	24.4	0.5
2000	15.8	0.8
1000	11.4	1.1
500	8.2	1.6
200	5.3	2.6
100	3.8	3.8
50	2.7	5.5
20	1.8	8.8
10	1.3	12.7

(ｲ)詳細集計

年平均結果の標準誤差

推定値の大きさ（万人）	標準誤差（万人）	標準誤差率（%）
5000	21.5	0.4
2000	13.0	0.7
1000	8.9	0.9
500	6.1	1.2
200	3.7	1.8
100	2.5	2.5
50	1.7	3.4
20	1.0	5.2
10	0.7	7.1

四半期平均結果の標準誤差

推定値の大きさ（万人）	標準誤差（万人）	標準誤差率（%）
5000	43.4	0.9
2000	26.0	1.3
1000	17.7	1.8
500	12.0	2.4
200	7.2	3.6
100	4.9	4.9
50	3.3	6.7
20	2.0	10.0
10	1.4	13.6

資料：平成21年　労働力調査年報

イ　地域別結果の推定値の大きさ別標準誤差

全国結果と同じ方法によって，結果数値の大きさ別に平均的に評価した地域別四半期結果の標準誤差率は，次表のとおりである。

(ｱ)基本集計

年平均結果の標準誤差率

推定値の大きさ(万人)	標準誤差率（%）									
	北海道	東北	南関東	北関東・甲信	北陸	東海	近畿	中国	四国	九州・沖縄
2000			0.3							
1000			0.5			0.4	0.4			0.4
500	0.6	0.5	0.7	0.4		0.6	0.6	0.5		0.6
200	1.0	0.9	1.2	0.7	0.6	1.0	1.0	0.9	0.6	1.0
100	1.5	1.3	1.7	1.1	0.9	1.4	1.5	1.3	1.0	1.4
50	2.1	1.8	2.4	1.6	1.3	2.1	2.2	1.8	1.4	2.0
20	3.4	3.0	3.9	2.8	2.3	3.4	3.6	2.9	2.4	3.2
10	4.9	4.4	5.7	4.1	3.4	5.0	5.3	4.2	3.6	4.5

四半期平均の標準誤差率

推定値の大きさ(万人)	標準誤差率（%）									
	北海道	東北	南関東	北関東・甲信	北陸	東海	近畿	中国	四国	九州・沖縄
2000			0.6							
1000			0.8			0.7	0.8			0.7
500	0.9	0.9	1.2	0.9		1.0	1.2	0.9		1.0
200	1.5	1.5	2.0	1.4	1.1	1.7	1.9	1.4	1.1	1.7
100	2.3	2.2	2.9	2.1	1.6	2.5	2.8	2.1	1.7	2.4
50	3.3	3.2	4.3	3.1	2.4	3.6	4.1	3.1	2.5	3.5
20	5.5	5.2	7.0	5.2	4.2	5.9	6.6	5.1	4.3	5.6
10	8.0	7.7	10.1	7.7	6.3	8.6	9.5	7.5	6.4	8.0

(ｲ)詳細集計

主要項目の年平均結果の標準誤差率

		北海道	東北	南関東	北関東・甲信	北陸	東海	近畿	中国	四国	九州・沖縄
役員を除く雇用者	標準誤差率（%）	2.3	1.6	1.3	1.5	1.8	1.4	1.5	1.8	2.0	1.5
	平成21年平均結果（万人）	(216)	(359)	(1501)	(403)	(226)	(629)	(795)	(297)	(144)	(543)
正規の職員・従業員	標準誤差率（%）	2.6	1.9	1.4	1.7	2.0	1.6	1.7	2.0	2.3	1.7
	平成21年平均結果（万人）	(137)	(245)	(980)	(270)	(160)	(421)	(515)	(202)	(99)	(360)
非正規の職員・従業員	標準誤差率（%）	3.2	2.4	1.8	2.1	2.7	1.9	2.2	2.6	3.1	2.0
	平成21年平均結果（万人）	(79)	(114)	(521)	(133)	(66)	(209)	(280)	(96)	(45)	(182)

資料：平成21年　労働力調査年報

(2) 非標本誤差

非標本誤差とは，誤差の要因のうち標本抽出（偶然性）に起因するものを除いた全ての要因により生ずる誤差をいう。それは更に，その要因により幾つかに分けることができる。報告者が質問を誤解したり懸念したりして事実と異なる回答をする場合の誤りや，無回答，調査員の面接の拙さによる誤り，不慣れによる標本の脱落・把握誤り，連絡・指導の不徹底による誤り，調査票の処理及び集計上の誤りなど幾らでも細かく分けることができる。このように，非標本誤差は調査のあらゆる段階で発生する可能性がある。

非標本誤差の特徴は，標本誤差のそれとは対照的である。すなわち，標本誤差が標本の大きさと密接な関係があり，避けられないものであること，量的な測定ができ，そのコントロールができることに対して，非標本誤差は標本の大きさと直接関係がなく，原因を究明すれば避けられるものもあること，量的な測定が難しく，そのコントロールができないことなどである。

調査が大規模になり関係者の数が増えると,非標本誤差の発生源も増えるものである。調査の各段階で誤りをできるだけ少なくし，非標本誤差を小さく抑えるには,調査関係者の努力と回答者の統計に対する理解に最も大きく懸かっている。

５　季節調整値

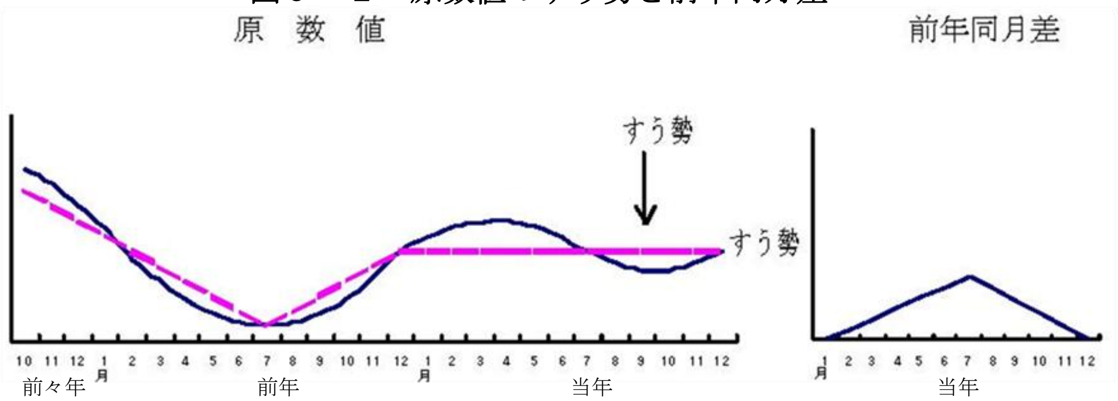
労働力調査の結果をみる場合，本章の４で述べた誤差の存在とともに注意する必要があるのは季節性の存在である。例えば，農業就業者は夏多く冬少ないが，これは毎年ほぼ決まって繰り返されるパターンであって，農林業就業者数のすう勢とは関係ない。したがって各月のデータを時系列的にみて農林業就業者が増加する傾向にあるのか否かを判断しようとする場合，季節的な変動による部分は除去して考えることが必要となる。

(1) 前年同月差の利用

前年同月差（又は前年同月比）を見るというのは，季節性を除いてデータを見る最も手軽な方法である。同じ月同士を比較すれば季節的な変動は自動的に除かれる。このため労働力調査の結果は，前年同月差又は前年同月比の形にして動向をみることが多い。労働力調査の標本設計においても，標本のうち半分を前年同月との継続標本とするなど，前年同月との比較の精度が向上するような工夫がなされている。

前年同月比較を行う場合幾つか注意すべき点がある。例えば，１年前との比較であるため，前年同月差（又は前年同月比）の変化方向は，基本的なすう勢の変化方向と必ずしも一致していない。前年同月差が拡大したといっても，それは１年前の動きによって引き起こされているかもしれない。例えば，図９－２のような例の場合，当年の前年同月差の動きは前年におけるすう勢的変化によって引き起こされたものであって，当年における実数値の動きは季節性を除けば安定したものとなっている。また，季節パターンが年とともに変化していく場合は，前年同月との比較では十分でないことに注意する必要がある。

図９－２　原数値のすう勢と前年同月差



(2) 季節調整の考え方

もう一つの一般的な方法として，季節パターンを除去する手法を原数値に適用して季節調整値を得る方法がある。季節調整値が得られれば，前月との直接比較が可能となる。

まず，原数値の動きが，次の四つの要素から構成されていると仮定する。

○すう勢変動（Ｔ：Trend）：経済の成長などに伴い，長期的に上昇・下降を示す変動

○循環変動（Ｃ：Cycle）：景気の循環に伴う変動など，ほぼ一定の周期を持つ周期変動で，周期が12か月を超えるもの

○季節変動（Ｓ:Seasonal）：12か月を周期とする変動

○不規則変動（Ｉ:Irregular）：上の三つ以外の変動で，突発的な出来事や，標本誤差などの変動

そして，原数値の系列Ｏ（Original）は

Ｏ　＝　Ｔ×Ｃ×Ｓ×Ｉ

という形で，各要素が乗法的に結び付いたものと仮定する。原数値の系列が，絶対水準が高くなるにしたがって季節的な変動の振幅も大きくなっているような場合は，このように仮定するのが一般的で，労働力調査の結果の季節調整もこの「乗法モデル」を用いている。

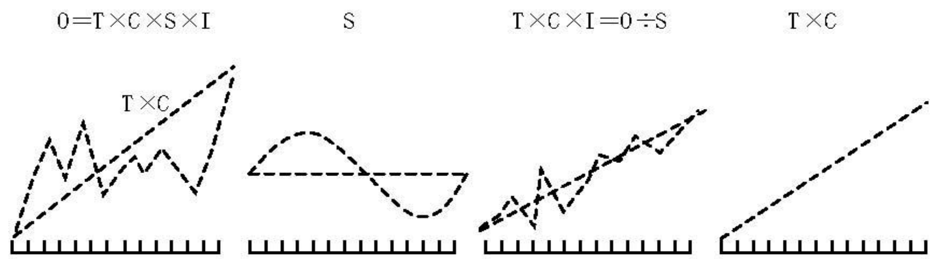
このとき，次の(3)で述べる方法などにより，年の各月のＳが得られたとする。Ｓはその年の季節パターンを示し，季節指数と呼ばれる。季節調整値は，各月において原数値を季節指数で除することによって得られる。式で示せば，

Ｏ÷Ｓ＝(Ｔ×Ｃ×Ｓ×Ｉ)÷Ｓ＝Ｔ×Ｃ×Ｉ

となる。このようにして得られた系列は，ＴＣＩ系列ともいう。

これを図で示せば，次のようになる。

図９－３　季節調整の概要



さらに，Ｉを除去すれば，すう勢変動と循環変動のみが残り，傾向的な動きを知ることができる。この系列をＴＣ系列ともいう。

(3) 季節調整の方法

季節調整を行う方法は幾つかあるが，一般的な方法としては，アメリカ商務省センサス局が開発したセンサス局法[注)]がある。センサス局法は，移動平均比率法に基づくもので，移動平均を繰り返すことによって季節変動成分を取り出そうというものである。センサス局法による基本的な季節調整の考え方は以下

注）センサス局は，1950 年代から研究を始め，1954 年にセンサス局法Ｉを完成し，その後複数次にわたる改正を行い，1965 年にセンサス局法Ⅱ(X－11)を公表している。センサス局法Ⅱ(X－11)は，多くのオプションを持ち汎用性に富むが，最近では，X－11 を改良した X－12－ARIMA を使用している統計も増えている。

のとおりである。

① 原数値の系列Ｏ＝Ｔ×Ｃ×Ｓ×Ｉ $\xrightarrow[\text{12か月移動平均}]{}$ $(\text{Ｔ×Ｃ})^{*}$

② Ｏ÷$(\text{Ｔ×Ｃ})^{*}$ $\longrightarrow$ Ｓ×Ｉ*

③ Ｓ×Ｉ* $\xrightarrow[\text{移動平均}]{}$ Ｓ

④ Ｏ÷Ｓ $\longrightarrow$ Ｔ×Ｃ×Ｉ　（ＴＣＩ系列）

⑤ Ｔ×Ｃ×Ｉ $\xrightarrow[\text{移動平均}]{}$ Ｔ×Ｃ　（ＴＣ系列）

※「*」は暫定値であることを示す。

実際には，このような手続を１回行っただけでは純粋な季節変動成分を安定的に取り出すことはできないので，⑤からまた②に戻るといったように，手続を何度か繰り返すようになっている。また，移動平均も様々な種類のものが使われている。

さらに，Ｏには，大規模なストライキや，天候不順等による例外的な変動が含まれていることがあり，その影響で季節変動成分，すう勢・循環変動成分の推定値がゆがんでしまう場合がある。センサス局法では，各成分を正しく推定するため，例外的な変動の部分を特異項として取り出し，それらを修正，あるいは除去して季節調整を行うようにしている。その場合，特異項と判定するためには，ベースとなるＴＣ系列が安定したものである必要があり，逆に安定したＴＣ系列を得るためには，特異項を修正しておく必要がある。このため，上に示した①～⑤のステップを，特異項を検出修正しつつ何度か繰り返すといった方法を採っている。

なお，センサス局法による移動平均の仕方，特異項の修正法については，「付録７　センサス局法の概要」で解説している。

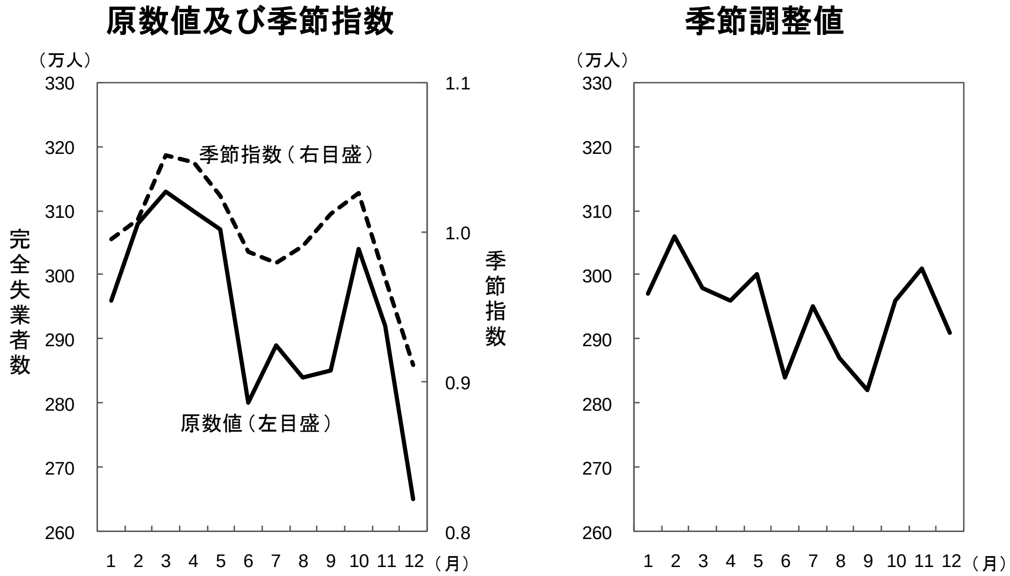
(4) 労働力調査結果の季節調整値

労働力調査では，センサス局法（X－12－ARIMA）のX－11デフォルト[注１)]による季節調整値を公表しており，現在約150系列[注２)]について，季節調整済みの月次又は四半期データがある。完全失業者数の季節調整値の例を示せば，図９－４のとおりである。

注１）特異項の管理限界は，下限 9.8σ，上限 9.9σとしており，これ以外は標準オプションとしている。

注２）開始年は系列により異なる。なお，最も長い系列は昭和 28 年１月からの月次データがある。

図９－４　季節調整値の例（完全失業者数－平成 17 年）



季節調整値をみる場合，二つの点に注意する必要がある。第一は，過去に公表された季節指数及び季節調整値は，年１回改定されるという点である。センサス局法では，過去の傾向から，将来の季節指数の推定値(推計季節指数)も計算できる。労働力調査では，例えば当年12月までの月次データがそろうと，それを使って翌年１月から12月までの推計季節指数を計算し，その推計季節指数により翌年の各月の季節調整値を公表している。そして，当年12月までのデータがそろうと，当年12月までのデータに基づき翌年の推計季節指数を計算するとともに，過去に遡って各月の季節指数及び季節調整値の再計算[注)]を行っている。このように，新たなデータが加わることにより，過去に公表した数値が改定されることに注意する必要がある。

第二は，各系列に対して独立に季節調整しているので，例えば男性の系列の季節調整値と女性の系列の季節調整値が，男女計の季節調整値に一致しないこともあり得る点（不加法性）である。また，完全失業率も，独立した一つの系列として季節調整している。このため，完全失業者数の季節調整値を労働力人口の季節調整値で除したものとは一致しない場合があることにも注意する必要がある。

注）原則として，当年から 29 年前までの原数値を用いて再計算を行い，直近 10 年分の結果について改定している。

６　時系列回帰モデルによる都道府県別結果の推定

労働力調査では，都道府県別に標本設計をしておらず（北海道及び沖縄県を除く。）[注1]，都道府県ごとの標本数も少ないことから，当初は都道府県別結果を集計していなかった。しかし，平成 13 年中頃に完全失業率（季節調整値）が調査開始以来初めて５％に達し，その後も高水準で推移し，厳しい雇用情勢が続いた。このような状況の中，雇用・失業情勢の詳細な把握に必要であることから，各方面から都道府県別の結果の公表が要請されるようになった。

このような背景から，都道府県別結果について，平成18年５月から時系列回帰モデルによる推定手法を採用し，より安定的な結果が得られるようにした上で，四半期平均結果（モデル推計値）の公表を開始[注2]した。

(1) 推定方法

労働力調査の都道府県別結果を推定する方法については，以下のような五つの要素からなる時系列回帰モデルを採用している。

$$Y(t) = X(t)\beta(t) + T(t) + S(t) + I(t) + e(t)$$

観測値　回帰　トレンド　季節変動　不規則変動　標本誤差

※観測値とは全国等の結果を求める方法（比推定）による都道府県別結果である。

それぞれの要素は次のような変動を表している。

○回帰項：各都道府県の動きと都道府県が属する地域（10区分）のトレンドとの関係を表す。

○トレンド項：経済の成長などに伴い長期的に変動を示すすう勢変動と，景気の循環に伴う変動などほぼ一定の周期を持つ変動で，周期が12か月を超える循環変動とを合わせた変動。景気の後退と回復によって，完全失業者が傾向的に増加したり，減少したりするような動きのことである。

○季節変動項：12か月を周期とする季節変動。

○不規則変動項：すう勢変動，循環変動，季節変動以外の変動で，突発的な出来事による変動や景気の短期的変動。地震などの自然災害や戦争などの国際紛争など一時的な現象の影響によって起こる生産の減少といった動きのことである。

注１）第８章の２を参照

注２）平成 14 年から，参考として比推定による都道府県別の年平均結果（試算値）を公表していたが，モデル推計値の時系列データが十分に整備されたことに伴い，平成 19 年平均結果をもって廃止した。

〇標本誤差項：労働力調査は，当月調査世帯の半分が前月・前年同月にも調査世帯となるような標本設計となっている。したがって，標本誤差は自己相関を持つ（前月・前年同月の標本誤差が大きければ，当月の標本誤差も大きい）とみなすことが可能である。そこで，これを仮定した時系列モデルにより，標本誤差と考えられる変動パターンと変動幅を前後の時系列データから推定したものである。

回帰項は，トレンドに近い変動を捉えており，回帰項とトレンド項とですう勢変動及び循環変動を合わせた変動と考えることも可能である。回帰項により，時系列的な変動要素に空間（地域）情報も取り入れることになり，より多面的な情報を推定に利用できるものになっている。

この推定方法による都道府県別の推計値は，比推定値（全国結果等と同様の推定方法）からカルマンフィルタ[注1)]を用いて，回帰項，トレンド項，季節変動項，不規則変動項及び標本誤差項に分解できる。こうして得られた標本誤差$e(t)$を従来の手法による推計値$Y(t)$から除いた値「$Y(t)-e(t)$」が，新たなモデルによる推計値[注2)]となる。

なお，相対的に標本規模の大きい北海道，東京都，神奈川県，愛知県，大阪府及び沖縄県については，比推定による推計を用いている。

注１）時系列データ（観測値）から誤差（ノイズ）を除き，真の値を推定する手法。航空機の飛行制御や GPS における計測データの処理（衛星配置や気象条件等により生じる誤差を除く。）など様々な分野で広く利用されている。

注２）毎年１～３月期平均の公表時に，新たな結果を追加して再計算を行い，前年までの四半期平均及び年平均結果を遡って改定している。