#### 研究成果報告書 科学研究費助成事業

今和 5 年 6 月 1 5 日現在

機関番号: 13101

研究種目: 基盤研究(C)(一般)

研究期間: 2020~2022

課題番号: 20K06276

研究課題名(和文)農業生産指標の標準設定の手法開発と実態解明

研究課題名(英文)Method development and actual situation elucidation of standard setting of agricultural production indices

研究代表者

平泉 光一(HIRAIZUMI, Koichi)

新潟大学・自然科学系(農学部)・准教授

研究者番号:00282997

交付決定額(研究期間全体):(直接経費) 2.000,000円

研究成果の概要(和文):農業生産管理や農産物原価管理等において、作物収量等の生産指標の適切な標準値の設定は避けて通れない。しかし、標準値の設定自体がないか、それがあっても適切ではない状況がある。本研究では、「典型値」や「正常値」とみなされてきた最頻値をもって、標準値とみなす立場を選択した。最頻値は、通例、ヒストグラムから求められるが、この方法での最頻値には端点の設定が任意になって一意性がない。そこで、一意的な最頻値として「鍾形頻度最頻値」と「代替最適平均」を考案して、標準値設定方法を提案した。

研究成果の学術的意義や社会的意義 農業生産管理や農産物原価管理等の農業経営管理においては、労働時間や作物収量等の生産指標の適切な標準値 の設定を行う必要があるが、営農の現場で標準値の設定が適切にできていないのが現状である。本研究では、最 頻値をもって、標準値とみなす立場を選択した。最頻値は、通例のヒストグラムから求める方法では一意性がな い。そこで、一意的な最頻値として「鍾形頻度最頻値」と「代替最適平均」を学術上新規に考案して、実際の農 業経営管理の改善に寄与しようとした。

研究成果の概要(英文): In agricultural production management and agricultural product cost management, it is inevitable to set appropriate standard values for crop yields and other production indicators. However, there are situations where standard values themselves are not set, or even if they are, they are not appropriate. In this study, we chose to consider mode, which have been regarded as "typical" or "normal" values, as standard values. Although mode is usually obtained from histograms, mode in this method is not unique because the anchor position is arbitrarily set. Therefore, "bell-form frequency mode" and "alternate optimum mean" were developed as unique mode, and the standard value setting method was improved.

研究分野: 農業経営管理(Farm Management)

キーワード: 農業生産指標 標準値 代表値 最頻値 鍾形頻度最頻値 代替最適平均 農業経営管理

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等に ついては、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

### 1.研究開始当初の背景

「農林業センサス」によれば、農家や法人組織等を合わせた農業経営体数は、2015年には137万7千経営体となり、そのうち組織経営体数は3万3千経営体となっている。農業経営体のうち農業サービス事業体等を含まない販売目的の法人経営体数(1戸1法人を除く)は約1万9千経営体で2005年の約2倍になり、法人化が進展している。1戸1法人を除く法人経営体を企業形態別にみると、株式会社等の会社の占める割合が64.2%と高く、次いで農事組合法人の占める割合27.4%となっている。このように、近年の日本の農業においては、大宗を占める家族経営が減少している一方で、雇用経営でもある企業経営が急増している。

農業経営管理は、企業経営になって、生産管理や財務管理も総じて家族経営よりも本格化・高 度化しているとはいえよう。だが、見聞する限りでは、金融機関や税務署などに開示する情報の 収集と活用である財務会計を最低限おこなってはいるものの、経営者自身が自らの経営のため に利用するものである管理会計は本格的に取り組まれていないようである。たとえば、利益を知 るために実際原価が計算されることはあっても、本格的な原価管理が行われることは稀である。 ここで原価管理というのは、予定原価の一種で科学的に決められる標準原価を事前に設定して、 経営の結果から分かる実際原価と比較して目標到達が出来たかどうかを判断して、目標到達が 出来ないときにはしかるべき手段を講じる一連のサイクルのことである。原価管理も広い意味 では、PDCAサイクル(計画: Plan 実行: Do 点検: Check 対応: Action)の一種である 経営管理手法である。農業経営において原価管理が浸透していない背景には、会社経営自体がま だ小規模な揺籃期にあって専門的な管理部門のスタッフが置けずに初歩的な経営管理しか手が 回らないということもあるが、商工業と比べて農業においては、計画時点で技術的な生産管理指 標の標準値を把握しにくく、その結果として標準原価も設定しにくいという特別な事情もある とみられる。例えば、耕種経営の作物収量については、生物学的に複雑な要因が影響し、また、 気象条件の影響で変動がみられ、事前に標準値が設定しにくいということが指摘できる。作業条 件の標準化と作業時間の標準値の設定は、学説的には近代経営学の始祖の一人とされるテイラ -が米国で 20 世紀前半に始めたものであり、今日では工場管理では必須のものとなっている。 農業経営にあっても、原価管理の前提として、標準時間や標準収量の設定は、本来必要なもので ある。ちなみに、標準時間とは、『生産管埋用語辞典』(日本規格協会、2002年)によれば、「そ の仕事に適性を持ち、習熟した作業者が、所定の作業条件のもとで、必要な余裕を持ち、正常な 作業ペースによって仕事を遂行するために必要とされる時間」である。標準収量については、こ れまで定義されている文献はみつからないが、習熟した農業者が、所定の土地条件・栽培条件の もとで、正常な気象条件のときに、適切な栽培管理によって達成される偶然的変動を捨象した収 量とする。なお、標準値は、当該経営にとって目標であると同時に予想の範囲内でもあるという 二重の性質をもつ。

我が国の原価計算論の第一人者である岡本清の代表的な著作『原価計算』(五訂版、国元書房、1994年)では、「標準作業時間の設定は、動作研究および時間研究を担当する技術専門家が行なう」とされている。大きな工場では時間技師の職能をもたせた専門的なスタッフが置けるが、殆どが小規模経営の農業経営ではそういう余裕はない。標準時間は、作業日誌をベースに当該経営内部で経営者が中心になって設定するしかない。標準収量も同様である。標準時間や標準収量は、農産物の標準原価の設定では必要なものであるが、これまでその水準を具体的にどう決めるかについては殆ど議論されてこなかった。阿部亮司「農業経営における標準原価計算」(農業計算学研究、1977年)は、農産物の標準原価計算に対する示唆に富む論考であるが、標準化の方法・根拠については触れられていない。北畠晶子「標準原価計算による溶液栽培作物の経営分析:コマツナを対象に」(農業経営研究、2002年)では、特定の農産物の実証的な原価計算を行っているが、標準原価の設定において「材料の消費量や生育日数、栽植密度、養液使用量などは標準的な値が示されていると思われる養液栽培システムメーカーのマニュアルを算出根拠とした」としており、また、「標準作業時間は過去の記録から能率の良い時の値を取った」としていて、標準設定としては根拠が弱い。以上から、農業生産指標の標準値を設定する実用的な方法を農業経営管理研究において取り組む必要がある。

# 2.研究の目的

そこで、本研究は農作業時間や作物収量のような農業生産指標の標準値設定を統計学的アプローチから行うことにした。農業生産指標の過去のデータから、標準値を求める際は、各種の代表値が候補になる。平均値は外れ値の影響を受けやすいので、平均値は標準値にはできない。外れ値の影響を極力小さくしたいだけであれば、ロバスト推定量も代表値の候補となりうる。だが、農業生産指標は、多くの場合、母集団が正規分布のような対称分布でないと予想される。例えば、大豆の収量は天候によって平年より3割減になることは結構ありえても、3割増になることは殆どありえない。稲の収穫で倒伏があると作業時間は通常の2倍以上になることもある。つまり、片側に裾の厚い分布でのピークとなる典型値に対応した代表値を考えるのであれば、中央値も含めた既存のロバスト推定量を適用するのでなくて最頻値を採用するほうが望ましい。当該経営の農業生産指標の過去のデータから、典型値となる代表値をもって標準値とみなすことにす

る。もちろん、これは、大胆な単純化であり、本来的には標準設定はもっと複雑になるはずである。だが、栽培・飼養や作業の高度な専門家とは限らない農業経営体の経営者が標準原価計算を行うためには、正常な原価から外れない程度の精度をもつ範囲で実際に適用できる標準値設定が求められる。農業経営研究としては、農業経営の経営者からみて実用的に採用が可能な範囲での経営者自らが生産指標の標準値設定を行う方法に限定して研究を進める方略の採用が望ましいと考えられる。また、最頻値を標準値として採用すれば、標準値が目標であり予想でもあるという性質をも満たすことになる。作業時間であれば、通常、最頻値は平均値より小さく、最頻値は実現可能な目標値となりうるし、また、予想の範囲内でもある。

しかしながら、カテゴリカルデータでなく収量や時間のような連続量については既往の最頻値計算方法には問題点が少なくない。ヒストグラム作成から求められる通例の最頻値計算法は端点(最初の階級の開始位置)が任意になるので最頻値が一意的にきまらない。一部の研究者が考案したヒストグラムを使わないカーネル密度推定法等の最頻値計算法も実用的には問題を抱えている。意外なことに、本研究の開始段階では実用的な最頻値計算法は確立していなかった。本研究では、生産指標の標準値を設定すべく最頻値を計算するために、実用に耐えうるヒストグラムによらない一意性をもった新たな最頻値の計算法を開発することを第 1 の課題とした。第 2 の課題としては、第 1 の課題の成果として計算される最頻値を参照基準としつつ営農現場における収量等の計画指標(可能であれば標準値)の水準設定の実態とPDCAサイクルの適用状況を解明することにした。ただ、コロナ禍で各地の経営体を訪問しての実態調査が出来なかったことから、第 2 の課題は実行できなかった。さらに第 1 の課題の補足的な課題として、最頻値が生産指標の標準値として適切な代表値であるかどうかも検討した。

# 3.研究の方法

日本工業規格 (JIS Z 8101-1:1999)では、最頻値について定義されている。その最頻値は、一般に流布している概念であるところの、最多出現値でも、また、ヒストグラムにおける最大度数のビンの階級値でもない。それは「離散分布の場合は確率関数が、連続分布の場合は密度関数が、最大となる確率変数の値」と定義されている(日本規格協会、1999)。本稿でも、原則的には、この定義に従う。ただし、日本工業規格では、「多峰分布の場合は、それぞれの極大値を与える確率変数の値」と規定されている。だが、本稿では、一般的な最頻値の概念に整合させるため、多峰分布である場合は最大の極大値をもたらす確率変数の値のみを最頻値として扱うことにする。実際問題としては、特定の農業経営で標準値を求めるうえで扱うデータは、標本サイズが高々1桁か2桁のオーダーであり、通常は大きな山が一つになる単峰分布であるのが殆どのはずである。大抵は、農業生産指標は、正規分布を片方に偏らせた歪みのある非対称分布か、正規分布に近似した分布になる。よって、多峰分布を考えないとしても、例外的なデータ以外では問題にならないはずである。なお、ここでは、位置パラメータが確率変数を指すものとする。

ヒストグラムを使う通常の最頻値計算法では、最大度数のビンの階級値を最頻値とみなす。この計算法は、一定のビン幅wを事前に指定したときの i 番目の観測値 x i の位置によって頻度の値が 1 か 0 をとる矩形頻度要素関数 y i ( m ) を導入して、その和を最大化するときの位置パラメータmを最頻値とみなす方法でもある。位置パラメータmが独立変数であるときに矩形頻度要素関数は、

IF | x i - m | w/2, THEN y i (m) = 1, ELSE y i (m) = 0 (1) となる( $i=1,2,\cdots,n$ )。矩形と呼ぶ理由は、関数の値が0より大きな部分の形状が、頻度( $i=1,2,\cdots,n$ )。矩形と呼ぶ理由は、関数の値が $i=1,2,\cdots,n$ 0。短形頻度関数  $i=1,2,\cdots,n$ 0。短子になるからである。このとき、矩形頻度要素関数の和である矩形頻度関数  $i=1,2,\cdots,n$ 0。数理計画法の用語でいえば、位置パラメータmの最適解は代替最適になる。ヒストグラムを作成して最頻値を試行錯誤的に求める通常の最頻値の計算方法は、そもそも最適解が一定の数値でなくて区間(点でなく線分)としてしか求められない方法である。頻度の要素関数を矩形にせずに、ガウス関数等を仮定すれば、カーネル密度推定法やミーンシフト法になる。それらの方法では頻度の要素関数を特定化しなければならないので、仮定が増えることになる。

適正なビン幅については、これまでに様々な提案がなされており、多様な選択が可能である。ここでは、Excel のデフォルトのヒストグラムのビン幅にも採用されていて比較的ポピュラーである Scott の選択に従うことにする。Scott (1979) は、標本の標準偏差を S、標本のサイズを n とするとき、ヒストグラムのビン幅wを次式のように提案した。

$$W = (3.49 \text{ S}) / n^{(1/3)} \tag{2}$$

以下では、最頻値を求めるときは、ビン幅として(2)式を選択することにした。

ヒストグラムによらない新たな一意的な最頻値を求めるための方法として 2 つを考えた。一つ目は、最頻値となる新たな代表値として鍾形頻度関数を最大化する位置パラメータの最適解としての鍾形頻度最頻値(bell-form frequency mode)を考案した。鍾形頻度最頻値を求めるために OR (Operations Research)の手法である数理計画法の枠組みから求めるための混合整数非線形計画モデルを構築することにした。二つ目は、最も単純で仮定の少ない方法として、矩形頻度要素関数の和である矩形頻度関数が最大となる代替最適区間の平均値をもって最頻値を推定する方法を考案した。この方法によって求めた最頻値を代替最適平均(Alternate Optimum Mean: AOM)と呼ぶ。代替最適平均は、表計算ソフトウェアのワークシートを用いて求める。

# 4. 研究成果

## 1)鍾形頻度最頻值

一意的な最頻値を求めるために、矩形頻度要素関数に代えて、鍾形頻度要素関数の和を最大化するときの位置パラメータmを考えたい。鍾形頻度要素関数 z i (m) は、位置パラメータmが独立変数であり、与えられた幅を v として、正値の観測値 x i が与えられているとき、

幅 v のとり方によって最適解の位置パラメータが変動するので、v の幅をどう決めるかが問題となる。ここではヒストグラムを作成して最頻値を計算する通常の最頻値の計算方法と親和性を持たせるようにv を設定する。具体的には、矩形頻度要素関数を位置について積分した面積がw (縦  $1 \times$  横w) になるので、f (m; x i) の正の部分を位置について積分した値もwになるようにv を決めることにする。

鍾形頻度最頻値を求める数理計画モデルは、次のような混合整数非線形計画モデルになる。

```
Maximize z i

s.t.

z i = i (1 - (xi - m)^2)

m 0

z i 0

i = 0 or 1

= 16/(9w2)

i = 1,2,...,n (4)
```

zi は変数の積を含む非線形の関数であり、0 から 1 の範囲の値をとる。 i は「インデケータ変数(ウィリアムス 1995)となるバイナリ型の整変数であり、論理的スイッチの役割を担う。 1 -  $(xi - m)^2$  が負になった場合、 i は 0 となり、1 -  $(xi - m)^2$  が非負になった場合、 i は 1 となる。ここで、 は xi - m が xi + 3 w / 4 になったときに xi = 0 とすべく  $(xi - m)^2$  が 1 となるように調整した係数である。

統計学的視点から既存の通常の最頻値計算法との関係を整理すると、次のことがいえる。鍾形頻度最頻値を求める最頻値計算法は、ヒストグラムを用いる通常の最頻値計算法と違って、一部の例外的データを除いて、代替最適に陥らずに一意的に最頻値が計算できる。鍾形頻度最頻値では、鍾形部分の幅のとり方にはスコットの選択に基づく階級幅をベースに使う方法以外もありうるものの、階級幅の選択を除いては誰が計算しても同じ値の最頻値が求まる。鍾形頻度最頻値は一意的な計算結果をもたらす故に、ヒストグラムを用いる通常の最頻値計算法よりも格段に精確でかつ実用的である。ただし、数理計画法のソフトウェアとそれを運用する知識が必要なので実務者が直接に計算することは想定しにくい。

#### 2)代替最適平均

代替最適平均は、表計算ソフトウェアで近似推定することができる。投稿中の論文では、 Excel2019 (Microsoft Corp.)のワークシートを使って、計算する手順を示した。

供試データは、2005 年から 2014 年の 10 年間についての日本の新潟県新潟市におけるコシヒカリという米の品種の年次ごとの 1 等米比率の時系列データである(新潟市役所調べ) データを昇順にソートすると、{ 11.6, 63.3, 71.5, 73.5, 75.2, 75.9, 83.2, 84.9, 88.7, 90.4} になる(単位:%) 標本の最小値は明らかに外れ値である。夏場の高温が原因で米粒の白濁が起きて極端に品質の悪い年があったためである。

設定した位置パラメータは2単位ずつで、11から91まで設定した。ワークシート上で、設定した位置パラメータ毎に、対応する観測値における1か0の矩形頻度要素関数の値を出して、その合計である矩形頻度関数の値を表示させた。矩形頻度関数の最大値は、ExcelのMIN関数を使った。使っている関数の性質上、最大値は局所最適でなく大域的最適になる。矩形頻度関数の最

大値は9で代替最適になっているのは、73、75、77、79、81の5つの値であった。その平均であるセルの値の77が代替最適平均の近似推定値となる。このデータの平均値は71.82で、中央値は75.55であるので、代替最適平均の近似推定値はそれらより大きな値になっていた。

代替最適平均を用いる方法の利点は少なくとも以下の5点が挙げられる。

第1に、代替最適平均は一意的な最頻値である。ヒストグラムを使って最頻値を求める通常の方法は、端点(あるいは階級位置)の設定が任意になって、求めた最頻値に一意性がない。代替最適平均を使った最頻値計算法は、その問題を解決できる。

第 2 に、代替最適平均を用いる方法は、ビン幅の仮定を用いるだけで、非常にシンプルである。頻度の要素関数の分布に特定の仮定を用いるカーネル密度推定法、ミーンシフト法および鍾形頻度最頻値法と比べると、また、密集したコアの個数の算出に正規分布の仮定を置くコア中央値法と比べると、代替最適平均を用いる方法は仮定が少なく、計算技法として非常に恣意性が低い。

第3に、代替最適平均は、表計算ソフトウェアのみで計算できる。代替最適平均は数理計画法の枠組みで求められるとしても、数理計画法に対応した特別なソフトウェアは必要ない。数理計画問題を解けるソフトウェアはそもそも代替最適解があると対応できないので、表計算ソフトウェアはその単なる代用品でもない。代替最適平均を用いる一意的な最頻値の推定方法は、基礎的な研究の成果であるだけでなく、十分に実用的であると思われる。

第4に、表計算ソフトウェアによる代替最適平均の推定は、近似解を求めるだけで厳密解を求めたことにはならないものの、実用的には推定精度を十分に確保できると考えられる。推定精度を上げようとすると、離散的な位置パラメータを非常に多数(数百から一千以上)設定しなければならない。だが、矩形頻度要素関数の式とそれらの和である矩形頻度関数の式は、最初の一つのセルのみに絶対参照と相対参照を適切に組み合わせて入力すれば、他のセルには式を複写するだけでよい。代替最適平均の推定値の精度を上げても計算コストはあまり増えない。

第5に、最頻値として代替最適平均を求める方法は、農業生産指標の標準値を求めるために開発された経緯があるにせよ、適用の対象は農業経営管理の分野に限定されない。代替最適平均を適用する用途には広範な汎用性があるはずである。

# 3)最頻値の標準値としての適格性

最頻値が生産指標の標準値として適切な代表値であるかどうかを検討する課題に取り組むにあたり、作物収量の時系列データ(過去6年、過去8年、過去10年)から最頻値を算出して、平年収量との誤差を求め、他の代表値(平均値、刈込平均、中央値)の平年収量に対する誤差と比較した。最頻値が生産指標の標準値として適切な代表値であれば、過去の作物収量の最頻値は平年収量と近似すると予想される。この予想が成り立つかどうかをデータに基づいて検討した。最頻値についてはコア中央値を採用して表計算ソフトウェアで計算した。作物としては、水稲と大豆を取り上げた。誤差としては、誤差の偏りの指標として、誤差の期待値(平均値)を設定し、また、誤差の散らばりの指標として、誤差の絶対値の平均値と二乗平均平方根誤差との2つを設定した。誤差を比較した回数は2001年までの連続する30年分である(大豆の平年収量の公開は2001年で終了しているので、水稲もその年次に合わせた)。作物収量と平年収量のデータは農林水産省の「作物統計」(長期累年統計)による。

計算の結果は次の通りである。水稲では、過去 6 年、8 年、10 年のいずれでも、誤差の偏りは、平均値、刈込平均および中央値よりも最頻値のほうが小さかった。水稲での誤差の散らばりは、過去 6 年と過去 8 年では最頻値よりも中央値の方が小さかったが、過去 10 年では最頻値の方が中央値よりも小さかった。大豆では、過去 6 年、8 年、10 年のいずれでも、誤差の偏りは、平均値、刈込平均および中央値よりも最頻値のほうが小さかった。大豆での誤差の散らばりは、過去 6 年、8 年。10 年のいずれも最頻値の方が、平均値、刈込平均および中央値よりも小さかった。

考察としては、上記の結果がどこまで一般性をもつかが問題である。おそらく、時系列でみた作物収量に非対称性がある限り、最頻値による平年収量からの誤差が、他の代表値による平年収量からの誤差よりも大きくならないことが少なくないと推認される。よって、作物収量という生産指標に関しては、最頻値が生産指標の標準値として適切な代表値であると考えられる。さらに、作物収量に限らず労働時間などの他の生産指標でも、生産指標のデータに非対称性がある限り、最頻値が生産指標の標準値として適切な代表値になりうると推察される。

# 5 . 主な発表論文等

「雑誌論文 〕 計2件(うち査読付論文 2件/うち国際共著 0件/うちオープンアクセス 1件)

| 「日銀誌論文」 T214(つら宜読刊論文 214/つら国際共者 014/つらオープンアクセス 114) |                    |
|---|--------------------|
| 1.著者名         平泉光一                                  | 4.巻<br>55          |
| 2.論文標題<br>農作業時間の最頻値としてのコア中央値の考案                     | 5 . 発行年<br>2020年   |
| 3.雑誌名 農作業研究   | 6.最初と最後の頁<br>79~86 |
| 掲載論文のDOI(デジタルオブジェクト識別子)<br>10.4035/jsfwr.55.79      | 査読の有無<br>有         |
| オープンアクセス<br>オープンアクセスとしている(また、その予定である)               | 国際共著               |

| 1.著者名  | 4 . 巻     |
|--|-----------|
| 平泉光一   | 57        |
|  |           |
| 2.論文標題   | 5.発行年     |
| 農業生産指標の標準値設定のための鍾形頻度最頻値の考案 : OR手法としての混合整数非線形計画法の応用 | 2022年     |
|  |           |
| 3.雑誌名  | 6.最初と最後の頁 |
| 農作業研究  | 231-238   |
|  |           |
|  |           |
| 掲載論文のDOI(デジタルオブジェクト識別子)                            | 査読の有無     |
| なし   | 有         |
|  |           |
| オープンアクセス   | 国際共著      |
| オープンアクセスではない、又はオープンアクセスが困難                         | -         |

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

| 6 | . 研究組織                    |                       |    |
|---|---------------------------|-----------------------|----|
|   | 氏名<br>(ローマ字氏名)<br>(研究者番号) | 所属研究機関・部局・職<br>(機関番号) | 備考 |

# 7.科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

| 共同研究相手国 | 相手方研究機関 |
|---------|---------|
|---------|---------|