

米国の潮流から日本の鑑定評価の方向性を考える ―統計学の活用を中心に―
〈上〉 A. I. テキストブック最新刊*¹における統計学の位置づけ

不動産鑑定士／大阪経済大学大学院経営学研究科非常勤講師 堀田 勝己

※本稿は、(株)住宅新報社より発行の『月刊 不動産鑑定』2014年12月号に掲載された論文である。

まえがき

本誌2014年5月号から11月号までの全7回にわたり、大野喜久之輔神戸大学名誉教授による「A. I. テキストブック 14版から米国鑑定業務の展開方向を探る」と題する論考が連載された。米国 Appraisal Institute が発行する不動産鑑定評価に関する最も重要なテキストブックである The Appraisal of Real Estate の最新刊（第14版）における旧版からの改訂内容に着目し、その核心部分を指摘した上で、米国における今後の鑑定評価理論と実務の行く先を探る内容であった。その第1回において、まず第13版と第14版の章立ての対比を行うことにより、全体構成の大きな変化が指摘されている。具体的には、次の如くである。

第13版では全30章＋補論A、Bで構成されていたものが、第14版ではまずパートIからXに大きく分けられた上で、章立ては全部で35章＋補論A、B、Cで構成されている。大きくパート分けされたことにより、鑑定評価の各作業内容に即して関連する章がまとめられ、従前よりもテキストとしてわかりやすい構成となった。また、パートの配列をみると、実務における作業手順を踏まえたものとなっており、同書が、実際に鑑定実務の現場で、マニュアル的に使われることを意識したものといえよう。

最新刊における特筆すべき変化（進化）のひとつとして、鑑定評価における統計的分析（Statistical Analysis in Appraisal）と題された章が、パートV「データの分析」内の第14章に置かれており、旧版においてやや傍論的に第28章に収録されていたのと比較すると、鑑定評価の作業手順の中の1プロセスとして、しっかりと位置づけられた印象である。

この変化に着目し、本稿では鑑定評価における統計学の活用という点にスポットを当て、米国テキストの記述が示唆する内容と、我が国の今後の鑑定評価の進むべき方向性について、2回にわたって考えてみたい*²。

まず今回は、表題の如く米国テキスト最新刊において、統計学に関連する章にはどのようなことが書かれ、どこまでの学習が要求されているのかを概観する。次回〈下〉においては、今回の内容を踏まえた上で、現状における日本の鑑定評価の問題点と、今後の課題のほか、次代を担う不動産鑑定士の養成教育のあり方についても言及したい。

1. 最新刊における統計学に関する記述

(1) 全体のパート割り

*1 Appraisal Institute, The Appraisal of Real Estate, 14th ed., Appraisal Institute, 2013

*2 本稿執筆の機会と貴重な助言を大野名誉教授から頂いた。ここに改めて謝意を表したい。

まず、A. I. テキストブック最新刊（第 14 版）のパート割りについて見ておく。

- Part I 不動産とその鑑定評価（Real Estate and Its Appraisal）
- Part II 問題の確認（Identification of the Problem）
- Part III 作業範囲の決定（Scope of Work Determination）
- Part IV データの収集と不動産の記述（Data Collection and Property Description）
- Part V データの分析（Data Analysis）
- Part VI 土地の価値に関する意見（Land Value Opinion）
- Part VII 評価手法の適用（Application of the Approaches to Value）
- Part VIII 試算価額の調整と価額に関する最終意見（Reconciliation of the Value Indications and Final Opinion of Value）
- Part IX 決定した価額の報告（Report of Defined Value）
- Part X 特殊な鑑定評価業務（Appraisal Practice Specialties）
- Appendix A 専門業務と法（Professional Practice and Law）
- Appendix B 回帰分析と統計手法の適用（Regression Analysis and Statistical Applications）
- Appendix C 財務公式（Financial Formulas）

上記のようにパート 1～10＋補論 3（A～C）で構成されており、統計学に関連するものはパート V 「データの分析」と補論 B 「回帰分析と統計手法の適用」がある。

（2）パート V の記載内容

パート V 「データの分析」は、次の 3 章で構成されている。

- Chapter 14 鑑定評価における統計的分析（Statistical Analysis in Appraisal）
- Chapter 15 市場分析（Market Analysis）
- Chapter 16 最有効使用の分析（Highest and Best Use Analysis）

上記のうち、筆者が本稿で取り上げる中心となるのは特に第 14 章（Chapter 14）であるが、同じパートに納められている他 2 章も大量のデータを収集、分析した上で判断が行われるという意味において、同じ作業の延長線上にある。

市場分析や最有効使用の分析というプロセスは、単に地域に精通し、相場観を持ち、土着的な視点でのみ行われるものではなく、鑑定主体がそれらの把握や判断に至る根拠となったデータとその分析方法、分析結果を示した上で行われるべきものであるから、その過程で統計的手法が用いられることは必然と言えるからである。

（3）第 14 章の置かれた意図

第 14 章の冒頭では、最近鑑定評価において統計学の必要性が増していることが述べられている。それは、対象不動産のみならず、比較可能な他の不動産、より大きな市場についての様々なデータセットに関する分析の結果が、鑑定主体の行う市場分析、最有効使用分析、あるいは評価手法の適用に必要な情報を与え、鑑定評価の結論に多大な影響を及ぼすからである。無論、データの分析それ自体は、従来から既に鑑定評価のプロセスにおいて

大切なものであったことは間違いないが、より厳格な分析方法や数値の解釈がなされるためには、統計学というツールの力添えが必要となる。よって、統計学の基本的な知識や技術を身につけることは、21世紀におけるプロフェッショナルと呼ばれるための前提条件となった、とまで同章では述べられている。

ことに回帰分析のような統計的手法は、取引事例比較法において対象不動産と比較対象となる取引事例との間の差異の把握のために有用な方法として受け入れられるようになった。歴史的に、大量評価などの局面では、その過程で統計的分析が行われてはきたものの、正式な統計用語が鑑定評価の文脈の中で用いられては来なかった。それだけ、両者の間には溝があったとも言える。

統計学は一般に、記述統計学 (Descriptive Statistics) と推測統計学 (Inferential Statistics) に分けられる。前者は要約尺度、図、表などを用いて標本 (サンプル) または母集団について記述するものであり、後者は標本 (サンプル) データを用いて母集団に関する推論を行うもので、母集団の平均値、中央値などの代表値、分散、因果関係などの基本構造に関する測定をも含む。

記述統計において表、チャート、グラフなどを効果的に活用するスキルはとても重要であり、鑑定評価書の中でこれらを用いて視覚化を図ることは、ユーザーの理解を促進し、信頼性を高めることに役立つ。但し、これらのツールを使いこなすためには、当然のことながら専門的な知識が必要とされる。統計学は、それ自体で独立した学問であるから、鑑定評価にこれを取り入れ、信頼性のあるサービスを提供するためには、どんな統計的手法を取り入れることが有効かについて、鑑定主体自身がよく理解していなければならないし、そのために必要な知識や技量を自ら習得しなければならない。

そのいわばスタートラインとして、第14章では、統計の基礎的な用語や概念を解説するとともに、鑑定評価における統計的分析の活用についても言及されているのである。

2. 第14章の内容

(1) データとは何か

データは、その性質により大きく2つのカテゴリーに分けられる。すなわち、質的 (Qualitative) データと量的 (Quantitative) データである。質的データはさらに、名義 (Nominal) 尺度と順序 (Ordinal) 尺度に分類され、量的データは間隔 (Interval) 尺度と比 (Ratio) 尺度に分類される。まとめると、[表1]のようになる。

質的データ (Qualitative Data)		量的データ (Quantitative Data)	
名義尺度 (Nominal Scale)	順序尺度 (Ordinal Scale)	間隔尺度 (Interval Scale)	比尺度 (Ratio Scale)

[表1]

質的データ

たとえば不動産の所在、地番、権利関係の記述などは、対象不動産を他と区別するため

に不可欠なデータであるが、定量的な意味を持つものではない。地番のような数字を含むものであっても、他と区別されるべき名前としての機能しか持っていない。これが、質的データの特徴である。

量的データ

上記に対し、地積、建物の床面積、築年数などはその数値自体に情報がある。これが、量的データの特徴である。量的データはいわば測定されたデータであり、平均値、中央値などの代表値で要約することもできる。

名義尺度

名義尺度は上述した不動産の所在、地番、権利関係の記述などのように、他と区別されるべき名前として定義されるだけであり、他との間に序列は存在しない。チューダー様式の建築がヴィクトリアンやコロニアルなどよりも上に位置づけられるといったことがないようである。

順序尺度

順序尺度も同様に名称として他と区別されるものではあるが、名義尺度とは異なり序列づけを行うことができるものである。状態の優劣をランクづけしたものは単なる分類ではなく、「優」は「劣」よりも上に位置することは明らかである。

間隔尺度

順序尺度と同様に、各データ間に序列があるが、異なるのはあるデータと他のデータとの間隔を測定することが可能な点である。ただし、間隔尺度には基準点やベースラインのようなものがない。2002年と2007年の間には5年の隔りがあり、言い換えれば加減することには意味があるが、乗除はできない。

比尺度

間隔尺度に対し、比尺度は加減のみならず乗除（比率）も可能である。例えばあるビルの賃貸可能総面積（Gross Rentable Area）が100,000平方フィートであり、使用可能面積（Usable Area）が90,000平方フィートであるとする、 $100,000 > 90,000$ という序列や、 $100,000 - 90,000 = 10,000$ という間隔だけでなく、 $100,000 / 90,000 \doteq 1.11$ という比率（賃貸可能 / 使用可能面積比率）にも意味がある。

（2）母集団と標本

母集団（Population）とは、あるカテゴリーに属するすべてのデータの集まりのことである。たとえば特定の地域に存在するガーデンレベル（半地下）アパートメントの全住戸といったデータが該当する。母集団の分布を記述するものを母数といい、平均、分散などがある。同地域に所在するガーデンレベル・アパートメントの平均床面積といった数値は、母集団の分布の様子を示す母数である。

これに対し、母集団から抽出された部分集合というべきものが標本（Sample）である。

標本の平均値から母集団の平均値を推定することは、推測統計の一例といえる。

標本に関する記述統計には標本サイズ、データの収集方法、収集時点などのほか、最小値、最大値、レンジ、四分位数、標準偏差、平均値、メジアン（中央値）、モード（最頻値）などによって分布の形状を示すことも含まれる。

標本データを用いて母集団パラメータを推定したり、母集団に関する一定の結論に到達するのが推測統計であることは既に述べたが、その活用の一例として、全米不動産協会（National Association of Realtors : NRA）が毎月公表している、合衆国全土の様々な市場における住宅価格の中央値に関する統計があげられる。このような統計により、その背後にある母集団の価格水準の変化といったものが推測されうるが、その信頼性や妥当性は、標本サイズや要因、標本が正しく母集団を代表しているかなどに依存する。

標本の抽出方法には確率抽出法（Probability Sampling）と非確率抽出法（Non-probability Sampling）があるが、標本の代表性を担保できるのは前者である。確率抽出法は、無作為抽出法（Random Sampling）とも呼ばれる。

上述の全米不動産協会による統計は、残念ながら非確率抽出法に基づくものであり、データの収集方法に基づく不確実性の度合いについてはわからない。

（3）分布の代表値

中央値

中央値（Median）とは、すべてのデータを最高値から最低値までその値の順に並べたときのちょうど中央の値である。もしデータ数 n が奇数であれば、中央値は $(n+1)/2$ 番目の値となり、 n が偶数であれば、 $n/2$ 番目の値と $(n/2)+1$ 番目の値の中間点となる。

たとえばデータ数が 15 の場合、中央値は 8 番目の値であるが、データ数が 14 の場合、中央値は 7 番目の値と 8 番目の値の中間点、すなわち 2 つの値を足して 2 で割った値となる。

分布に極端な値が含まれる場合、下記の算術平均等ではその影響を受けて分布の全体像をうまく表さない場合があるが、中央値ではそのようなことがない。

算術平均

算術平均（Arithmetic Mean）は、最もよく用いられる代表値であり、標本平均、母平均などとして用いられる。すべてのデータを合計し、その個数で割ったものである。分布に極端な値が含まれる場合、算術平均はその影響を受け、分布の姿を正しく表さず、代表値としてあまり適切ではない場合がある。

標本平均には \bar{x} の文字が用いられることが多く、母平均には μ の文字が用いられることが多い。

分布型が知られている場合、あるいは標本サイズが十分に大きい（標本数が十分に多い）場合には、標本平均は統計的推論にたいへん役立つものである。

$$\text{算術平均} \quad \bar{x}_A = \frac{\sum_{i=1}^n x_i}{n}$$

幾何平均

金融の世界における投資収益に関してよく用いられるのが、幾何平均 (Geometric Mean) である。データ数が n の場合、 n 個すべてのデータを掛け合わせた値の n 乗根をとった値のことである。

$$\text{幾何平均} \quad \bar{x}_G = \sqrt[n]{\prod_{i=1}^n x_i}$$

最頻値

最頻値 (Mode) とは、データの中で最も頻繁に出現する観測値のことである。最頻値が一意に定まらず 2 つあるデータ群の性状を二峰性 (Bimodal) といい、それより多いものを多峰性 (Multimodal) という。一部の分布型 (一様分布など) では最頻値は定義されない。

(4) 散布度の指標

標準偏差と分散

いずれもデータのばらつきを示す指標である。

一般に、母分散は σ^2 の記号で表し、標本分散は S^2 で表す。また、分散の平方根が標準偏差である。

母分散 (Population Variance) は、各観測値の平均値からの隔たりを二乗したものの平均であり、その平方根が母標準偏差 (Population Standard Deviation) である。

$$\sigma = \sqrt{\frac{\sum (X_i - \mu)^2}{N}} \quad N: \text{母集団の総観測数}$$

標本分散^{*3} (Sample Variance) は、各観測値の平均値からの隔たりを二乗したものを自由度 ($n-1$) で割ったものであり、その平方根が標本標準偏差 (Sample Standard Deviation) である。

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \quad n: \text{標本サイズ}$$

変動係数

変動係数 (Coefficient of Variation) は、標準偏差を平均値で標準化したもので、複数のデータセット間のばらつきの度合いの比較に有用である。

$$CV = \frac{S \times 100}{\bar{x}}$$

レンジ (分布幅)

データの幅のことであり、単純に最小の観測値と最大の観測値の隔たりのことである。

*3 母集団の分散の不偏推定量であり、不偏分散と呼ばれる。標本自体の分散ではないので、標本分散の語は誤解を招く表現だが、実際にはかなり多用されている。

四分位数と四分位範囲

データ群を値の順に並べ、中央値で二分する。下の組の中央値を第1四分位数 (Quartile1 : Q1) といい、上の組の中央値を第3四分位数 (Q3) という。定義から、第2四分位数 (Q2) は、データ群全体の中央値そのものである。

$$Q1 = \frac{n+1}{4}$$

$$Q2 = \text{median}$$

$$Q3 = \frac{3(n+1)}{4}$$

四分位範囲 (Interquartile Range : IQR) とは、第3四分位数と第1四分位数の差 (Q3-Q1) のことであり、中央値のまわりの50%のデータがこの範囲に含まれていることになる。

(5) 分布型に関する指標

分布型がどれだけ正規分布に近いのか、中央値と平均値の隔たりがどれほどあるかなどを知ることができる指標である。

歪度

歪度 (Skewness) は、分布のゆがみ、左右対称の度合いを示す。正規分布 (Normal Distribution) は中央値と平均値が同じで、左右の歪みがないので、歪度は0である。

尖度

尖度 (Kurtosis) は、分布の尖りの度合い、すなわち分布型の高さや裾の厚さを示す。正規分布は尖度=3 とされる*4。

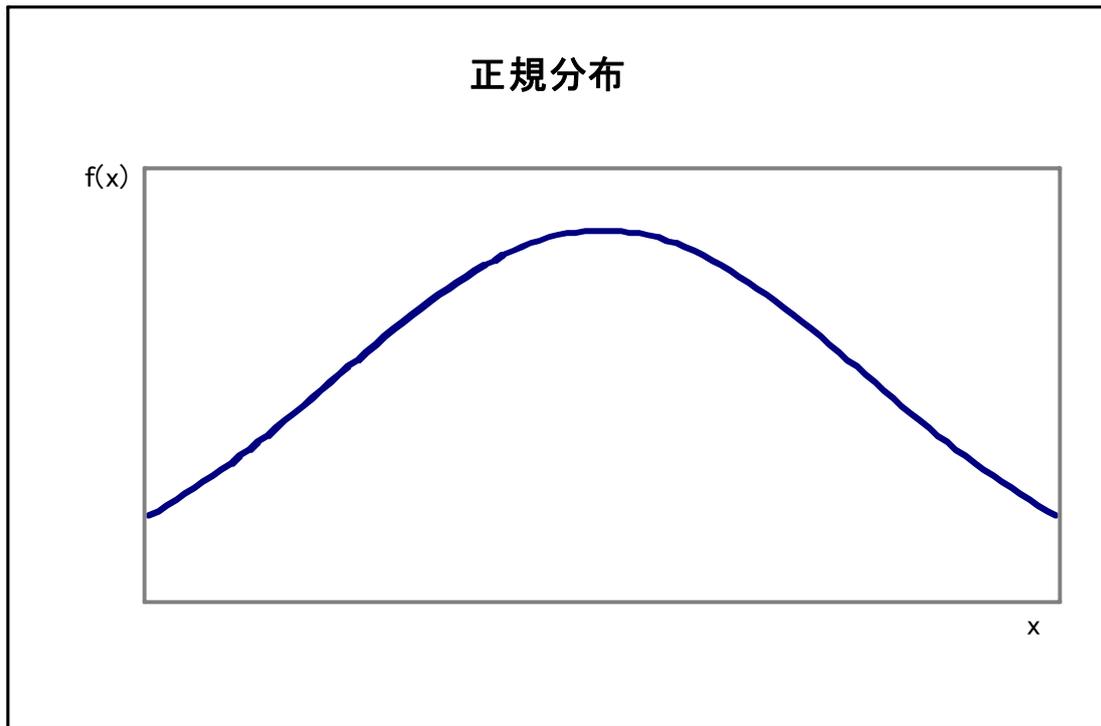
正規性

正規分布 (Normal Distribution) は[図1]のような左右対称のつりがね型 (Bell Shape) であり、平均値=中央値=最頻値である。また、上記のとおり歪度=0、尖度=3となる。

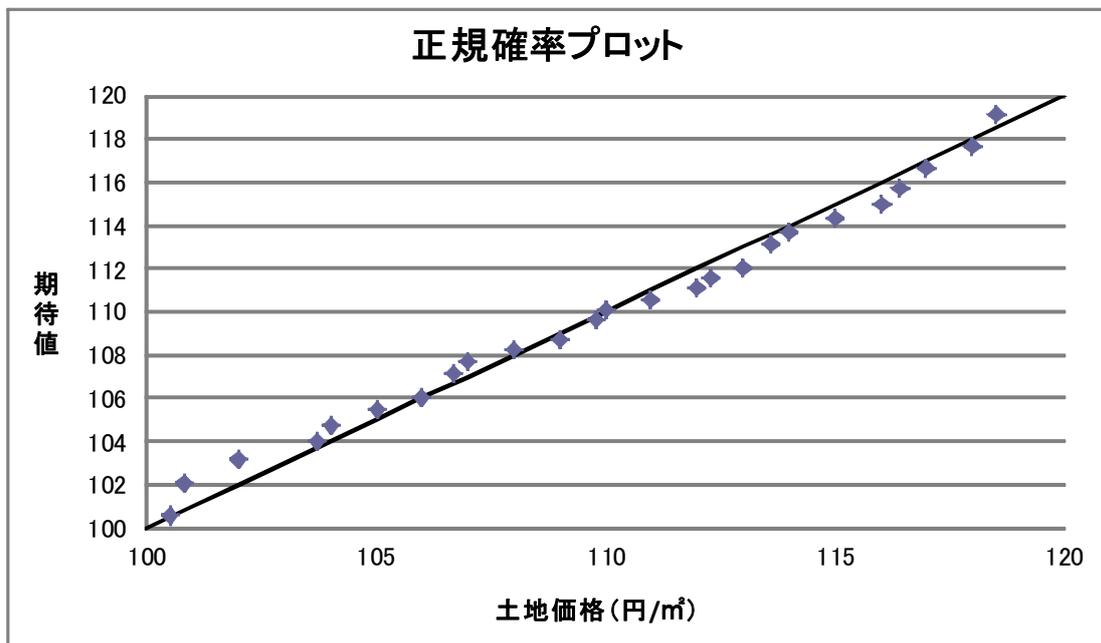
母集団が正規分布に従っているかどうかを調べるために、正規性の検定 (Test of Normality) が行われる。また、視覚的に確かめるために、正規確率プロット (Normal Probability Plot) が作成される。Minitab、SPSSなどの統計ソフトを用いると、簡単に行うことができるが、Microsoft Excelでも同様のことができる。

[図2]は、筆者が土地価格データ (想定) の標本 (30個) を用いて、Excelで正規確率プロットを作成した例である。正規分布に従う場合、斜めに引いた45度線にぴったりと合致し、ずれが大きいほど正規分布からは遠いことが視覚的に確かめられる。

*4 テキストでは正規分布の尖度は3とのみ書かれているが、0とする定義もある。



[図 1]



[図 2]

(6) 中心極限定理

統計的推論において、常に母集団が正規分布であることを前提とすれば分析は容易となるが、すべての現象が正規分布するわけではない。しかしながら、母集団がどんな形の分布であっても、そこから取り出した標本のサイズが十分に大きい場合には、標本の平均は

正規分布に近づく。これを中心極限定理 (Central Limit Theorem) という。これにより、実際には正規分布しない現象についても、正規分布の理論を用いて分析を行うことが可能となる。

母分散が未知で、標本から母平均の区間推定を行う際、信頼区間 (Confidence Interval) はスチューデントの t 分布 (Student's t distribution) によって求める。

$$\text{信頼区間: } \bar{x} \pm t_{n-1} \frac{S}{\sqrt{n}}$$

(7) 回帰分析

回帰分析 (Regression Analysis) は、従属変数 (Dependent Variable) と独立変数 (Independent Variable) との関係を定量化して関係式を導き出す統計的手法である。鑑定評価において、通常、従属変数は価格または賃料である。独立変数は、土地や建物の価格に影響を与える4つの要因 (社会的、経済的、行政的、環境的) と物理的特性などである。

回帰モデルは、従来から、資産税の徴税のための大量評価 (Mass Appraisal)、ことに発展した住宅市場などにおいて活用されてきた。なぜなら、一般の鑑定評価に比べて、効率的であり、資料の制約などにより個別の鑑定評価を行い得ないような場合にも有効だからである。

(8) 鑑定実務における統計的手法の活用

パーソナルコンピュータと表計算プログラム (Spreadsheet Program)、統計ソフトウェアなどの普及により、鑑定士は簡単かつ正確に、鑑定評価に統計的手法を取り入れることが可能となった。コンピュータがビジネスに導入された初期は、統計的手法といってもほぼ記述統計に限定されていたが、オペレーティングシステムに視覚的なインターフェースが導入されるとともに、SPSS、Minitab、SAS などの使いやすい統計ソフトウェアの出現により、ユーザーはもはや自らプログラミングをする必要がなくなった。

現在では、Microsoft Excel のような表計算ソフトウェアによっても相関分析、F-検定、t-検定、回帰分析などが行えるようになっている。

自動評価モデル (AVM)

今日、米国の鑑定士たちに利用されている自動評価モデル (Automated Valuation Models) は、資産税の徴税職員たちが効率性と公平性を求めて開発した大量評価のための技術に由来している。彼らは、直接コンピュータに読み取れる形式のデータを大量に入手できる特別な地位にあったために、それが可能となった。

AVM に組み込まれているニューラルネットワーク (Neural Networks) は、変数間の関係性を自ら「学習」し、価格算定のための未知のアルゴリズムを自動生成し、継続的に更新してゆく。ただ、ブラックボックスであるニューラルネットワークは、算出された結果においてのみその妥当性がテストされるために、“非論理的”とみられている。それゆえ、米国鑑定業務統一基準 (Uniform Standards of Professional Appraisal Practice : USPAP) においても、AVM から算出された結論それ自体は、鑑定評価ではないと明言されており、鑑定士がそのアウトプットが妥当なものであると判断する限りにおいて、自らの鑑定評価にお

ける結論づけの基礎として、利用することができるとしている*⁵。

現在では、AVM は全国的に提供され、次々と新製品も出現しており、金融業界など様々な業界において採用されている。

統計的手法を取り入れたカスタム評価モデルは、鑑定実務に新たな機会を提供するものといえる。商業市場に比べ、分析に必要なデータ量が得られやすい賃貸住宅市場などの評価において、十分な統計モデルの設定スキルを備えた鑑定士が適切なソフトウェアを用いることにより、特殊な評価ニーズにも対応することができる。ソフトウェアによっては、モデル構築などが簡単に行えるものと、複雑で難解なものがあるが、鑑定士は自らの能力や経験を越えたモデルを用いようとしてはならない。鑑定評価における様々な素養と同様に、統計的な問題解決能力も、経験によって培われる。統計分析の専門家などと連携し、教育の継続によって、その能力を構築してゆくことが肝要である。

3. 鑑定評価と統計学～不動産鑑定士に要求されるもの

以上、A. I. テキストブック最新刊第 14 章の記述内容を概観したわけであるが、統計学としては極めて初歩の基礎的知識を整理した程度の内容となっている。

補論 B において、「回帰分析と統計手法の適用」として線形単回帰 (Simple Linear Regression) と重回帰 (Multiple Linear Regression) について、具体例を挙げて解説されているが、その内容については、次回に譲ることとする。

不動産鑑定士にとって、統計学はあくまでもツールであり、それを極めることは目的ではないから、まずは基礎的な統計学の素養を身につけ、ソフトウェアの操作方法と、そこからアウトプットされる結論について十分に理解できるようになることが必要と考えられる。

A. I. テキストブック主要部分に 1 章を割いて統計分析に関する記述が掲載されたこと (既述のとおり、第 13 版においては終盤の傍論的位置づけであった) は画期的なことである。コンピュータが発達し、社会におけるデータベースの構築が進んだことによって、大量のデータを扱うことのできる統計的スキルを身につけることは、いまや米国のみならず、日本におけるどんな業界においても不可避のことといえる。「統計学が最強の学問である」*⁶と題する著書が、ビジネス書として大ヒットしたことは記憶に新しいが、それだけまだ多くの人が、統計学の本当の効用を知らなかったり、不用意に持ち上げたり、不当に批判をしたりしていることの裏返しでもある。

不動産鑑定士試験の経済学の出題をみると、近年、一部計算能力が問われるような問題が散見されるようになったが、当業界に必要とされる人材の条件として、数的な処理能力を備えることが求められているといえよう。試験合格者が受修する実務修習においても、統計学の講義は (以前から) 行われている。少なくとも「数式は苦手」と敬遠してはこれらに対応することは難しい。

次号では、テキストブック補論に掲載されている回帰分析の適用例について紹介し、鑑

*5 Advisory Opinion 18, Uniform Standards of Professional Appraisal Practice: Use of an Automated Valuation Model (AVM).

*6 西内啓「統計学が最強の学問である」ダイヤモンド社, 2013

■不動産鑑定士堀田勝己の WEB SITE■
<http://www.kanteishi.net/>

定評価においてこのような分析を行う意味について考えるとともに、日本の鑑定評価の現状における問題点や、今後の展開についても言及したい。

以上