

統計ずさん処理問題を考える

関西大学 社会安全学部 小澤 守

先ごろの国政における基幹統計のずさん処理の報道をみて考えた。そもそも統計では、1. 何のために統計をとるのか、2. データの収集方法はどうか、3. データの信頼性はどうか、4. 集めたデータの処理方法はどうか、などが問題となる。今回最も大きな問題となった厚労省では検証さえも問題を引き起こし、そのために政策統括官を更迭し、予算委員会で苦しい答弁に終始している。一方の野党は、なんでこんなことが起こったのか、好景気感がない、統計資料が間違っているから実際の経済成長率はもっと低いなどと政府追及の小道具として活用するなど、にぎやかなことこの上ない。新聞報道によれば、総務省の調査で56ある基幹統計のうち不適切な処理があったのは厚労省のものにとどまらず22もの統計にデータの誤りや変更の未申請などの問題が発覚し、関係するのは国交省、総務省、経産省、財務省、文科省、農水省など極めて多数に及ぶという。これも新聞報道によれば、我が国政府の統計担当職員数は1,940人、米国は13,000人、カナダで5,000人と極端に少ない。そもそも政府の統計は全数検査を行うものから全数検査を行わないサンプリング調査までさまざまであるが、その結果が政策や予算、さらには家計にも影響するものであれば大変なことである。その中でもどうやら厚労省が一番深刻で、予算案の修正が必要な毎月勤労統計が対象で、結果として政策統括官の更迭や予算委員会招致が求められているのである。

このシリーズではすでに企業におけるデータ改ざん問題を取り上げたが、今回の問題もそれに匹敵する重要な問題であるのは間違いない。我が国は公的データの信頼性に大きな問題を持つ国として海外に知らしめた感があるが、予算委員会での政府側、野党側の議論を聞いていても、本稿の最初に書いた4つの論点を中心課題とせず、政府追及の単なる手段としての議論に終始しているようにも思うがいかがか。全数調査すべきところをサンプリングしたとか、補正したとかはまだいいとして、どんな補正なのか、得られたデータ処理方法は、また調査方法や調査対象の妥当性、つまりはデータの信頼性についての議論があまりなされていないように思う。

工業分野では統計的品質管理法などが従来から行われている。筆者が統計処理に関連して何度かお世話になる統計の本、G. W. Snedecor著 *Statistical Method* (邦訳は『統計的方法』として出版) には、「サンプルが母集団を正確に代表しているなら、統計の結果は妥当なものになる。最大限の正確な情報をどうやれば集められるのかについては十分な思考と長期にわたる経験が必要である。調査にあたった者の最良の努力なくしてはサンプリングの結果は母集団を代表する結果からほど遠いものになってしまう」と警告している。至極もつともで、サンプリング調査の基本中の基本である。上記の毎月勤労統計をはじめとする各種統計データの集め方が正当に国民の生活を代表する数値なのか、統計にあたるものは絶えず反芻し、様々な試行もしなければならないと思う。昔に決めた統計の取り方を後生大事にまた機械的に（この表現は機械屋としては気に入らないのだが。「無意識に」というのが

好ましい) 維持するのも問題だし、かといって予算の都合や人員不足を理由として勝手にデータサンプリング方法を変えられるのも問題である。統計の先進国である米国は上記のように13,000人の要員を抱え、人口当たりで言えば我が国の2倍程度の数となる。行政に統計を取り入れる手法はおそらく明治のころから、いや第二次世界大戦後に欧米特に米国から輸入したものであろう。根本的な手法など十分な検証もせずに機械的にやっているように思えてならない。専門官の不足を言うなら、政府も各省庁に統計部署を置くのではなく、各省に分散しているサンプリングから統計処理まで行う部署を総務省などに集約するのも方法であるが、統計データは本来的には行政の検証データでもあることから、民間の第三者的研究機関を組織してそこに集約すべきである。そうすれば統計の専門家の養成も、官僚のキャリアパスなど無関係に行える。統計データで飯を食い、その信頼性が揺るげば企業として成り立ちえないような組織でないと、喉元過ぎればのたとえで、いつまでたってもきっと同じことを繰り返すのではなかろうか。

統計データは担当省庁や部署固有のものではなく、国民全体のものであることを強く認識しなおすことが重要である。「統計とは何か」、国民等しく、当然政府省庁職員や議員も含めて、改めて考える必要があるようだ。

