

## 統計ずさん処理

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。そもそも統計処理は1. 何のため、2. データの収集方法は、3. 集めたデータの処理方法は、などが問題となる。今回最も大きな問題となり厚労省では政策統括官を更迭、野党は高景気観がない、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、予算委員会のにぎやかなこと。新聞報道によれば不適切な処理があったのは厚労省にとどまらず国交省、総務省、経産省、財務省、文科省、農水省など、極めて多数に及び、不適切の内容は必要な手続きを怠った単純ミスで、我が国は政府の統計担当職員数は1940人、米国は13000人、カナダで5000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、さらには筆者の家計にも影響する年金にも影響する非常に重要なものである。総務省の調査によれば56ある基幹統計のうち23もの統計にデータの誤りや変更の未申請なども問題があったとのことである。その中でもどうやら厚労省が一番大変で、予算案の修正が必要な毎月勤労統計だそうで、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。

先に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はないが、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3つの論点のいずれも、非常に分かりにくい。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らないように思う。今一度、統計の在り方について検討が必要に思うがいかがか。

工業分野では統計的品質管理法などが従来からあり、また筆者が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心にとめ、様々な試行も行わなければならないと思う。その昔、統計の取り方をいったん決めたから後生大事に同じ方法を機械的に行うのも問題だし、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13000人の要員を抱え、人口当たりで言えば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欧米の手法の輸入品で、我が国独自でもない。

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。政府も各省庁に統計部署を置くのではなく、総務省でもいいからすべての統計をサンプリングから処理まで行う部署に集約して、統計の専門家を養成すべきである。かつて交通事故による死者数が6000人だったか7000人だったか忘れたが、ある時から警察庁が交通事故から24時間以内に死亡した人数だけ数えるようになって、交通事故死の統計データがいきなり4000人台にまで落ちたのを記憶している。一見すると日本では交通事故死が急に減少したかのような錯覚を覚えたものである。長期にわたる事故件数などのグラフを書き続けていると大きな段差ができてしまうのである。それはそれで致し方ないのであるが、その数が例えば交通事故死に対する保険金額などの算定基準に使われるようになると、大きな問題を引き起こしてしまうのである。統計データは担当部署固有のものではなく、国民全体のものであることを改めて強く認識することが重要である。今一度、「統計とは何か」、国民等しく、当然政府や議員も含めて、考える必要があるようだ。