

論文の内容の要旨

論文題目 超母集団モデルによる寸法指標推測について
氏 名 星野伸明

ある集団が、多数の副集団からなるとしよう。ここで特定の個体数の副集団数を「頻度の頻度」または「寸法指標」と呼ぶ。統計的生態学、計量文献学、統計的開示制限等の分野では、母集団の寸法指標を標本から推測する必要性が有る。

古典的な有限母集団解析では、母集団の構造を仮定しないネイマン流の手法が用いられる。しかしネイマンの枠組みで寸法指標の唯一の不偏推定量は、存在しても精度が不十分である。従って母集団に関する構造を仮定して情報を補う事で、推定精度を確保するのが現実的な選択である。そして母集団構造を記述するには、超母集団モデルが便利である。

超母集団モデルでは、現実の母集団を仮想的な無限母集団（＝超母集団）から抽出されたものとみなす。つまり超母集団の仮定は、無限母集団分布として明示的に与えられる。このような手法を用いる場合、先見情報に適合的な挙動をする確率分布が興味の対象となる。経験的に寸法指標は右裾の長い非負整数上の分布で良く記述されるので、これをモデルに活かせば良い。

著者の方法論ではより広く考えて、モデル集合の設定が先見情報の定式化という事になる。つまり、一つのモデル・分布を事前に特定するのではなく、複数のモデルを集合的に候補とする。ひとたびモデル集合を定めれば、その中からデータ所与でモデルを選択すれば良い。そして、選択されたモデルによる推測を採用する。このように考えた場合、使える超母集団モデルを多く用意する事が重要となる。何故なら多くの異なるモデルによって集合を構成すれば、良く記述できるデータが増える。そして、より尤もらしい推定が可能になる。ただし多くのモデルを扱うには、その体系的な理解が必要となろう。従って、寸法指標モデリングの一般論が考察される。

まず、超母集団モデルによる母集団寸法指標推測値の精度が検討される。利便性の観点から、単純な経験ベイズ法が前提になる。この場合点推定値は、モデルの母数の推定値の下での母集団寸法指標の期待値で良い。しかし区間推定をモデルの母数の推定値所与で構成すると、余りにも正確に

見えすぎる結果となる。つまりこのように構成された信頼区間は、母数の推定誤差を考慮に入れていない他、モデル選択における不確実性も反映されていない。もちろんこれ等の要因で信頼区間を補正する事も、原理的には可能である。しかし正確な区間推定が必要ならば、ブートストラップのような数値的方法を用いる方が便利であろう。

次に著者は、複合ポアソン分布について考察する。右裾が長い非負整数上の分布の多くは複合ポアソン分布なので、その性質を明らかにする事は重要である。特に、複合ポアソン分布の無限分解可能性が議論の中心となる。この性質に依存して、複合ポアソン分布の独立同一分布モデルに小数法則とも言うべき極限操作が適用出来る。そしてその極限では、各寸法指標が独立にポアソン分布に従うという簡潔なモデルが得られる。これらの結果と母集団サイズに関する条件付けを組み合わせれば、複合ポアソン分布によるモデリングは4形態に整理される。

なお小数法則の極限は、無限個の正の整数上の分布からなるモデルという解釈を持つ。従って寸法指標が独立にポアソン分布に従うようなモデルを用いれば、有限個の正の整数上の分布による伝統的なモデリングが伴う不可知量への依存という困難を回避できる。

以上の一般的議論を背景に、個別の超母集団モデルに関する文脈が整理される。具体的には各モデルの導出、他モデルとの関係、寸法指標のモメント、母数推定に関する結果が示される。取り上げられているモデルは、ディリクレ=多項、ガンマ=ポアソン、対数正規=ポアソン、一般化逆ガウシアン=ポアソン、逆ガウシアン=ポアソン、対数級数、Ewens、Pitman、条件付逆ガウシアン=ポアソン、拡張負の二項、極限条件付逆ガウシアン=ポアソンである。また、ポアソン・パスカルモデルについても考察される。

最後に、超母集団モデルによる寸法指標推測応用例が示される。統計的生態学、計量言語学、統計的開示制限の問題意識がそれぞれ紹介された後、現実のデータにモデルが適用される。