

R と Excel を用いた分布推定の実践例

近藤宏樹（日新火災海上保険株式会社）

斎藤新悟（九州大学マス・フォア・インダストリ研究所）

当発表の内容は、日新火災海上保険株式会社と九州大学との共同研究の成果である。

モーメント法や最尤法といった分布推定の実装は Excel でも可能だが、元データの数が多い場合は計算に時間がかかる。そこでこの発表ではまず、統計解析ソフト R を用いることで分布推定の正確化・高速化を図る一方で、推定結果の表示には表の視覚化に優れた Excel を用いた実践例を紹介する。次に、分布推定におけるパラメータリスク・モデルリスクの評価について、ある特別な状況で Bayes 推定を用いてリスクを計測する具体的な方法を提案し、R を用いた実装例を紹介する。以下、それぞれの内容についてもう少し詳しく説明する。

R と Excel を用いた分布推定

R はオープンソースでフリーな統計解析ソフトであり、世界中の研究者によって数多くの統計的手法が実装されている上に、グラフィックにも優れているという特徴を持つ。昨年度の年次大会で「損保のアクチュアリー業務における R の活用」と題したプレゼンテーションが行われたことに象徴されるように、今後保険業界においても R が普及することが期待される。

分布推定に関しては、分布形とパラメータの推定方法を指定すれば、パラメータの推定値に加えてフィッティングの良さを表す様々な統計量も計算するような R のパッケージが公開されている。この発表では分布推定自体にはこのパッケージを用いる一方で、分布推定の結果である推定値や統計量を一覧表示する際には表の視覚化に優れた Excel を用いる。

さらに、分布推定の結果から適切な分布形を選択する際には、このパッケージが自動的に計算する統計量のみを利用することも可能であるが、統計量によって適切であると判断される分布形が異なることもあり、最終的には種々の情報を加味して総合的に判断を下すべきであると考えられる。R は判断の材料となるいくつかの図を表示することができるため、ここではそれも合わせて判断するプロセスを紹介する。

Bayes 推定によるパラメータリスク・モデルリスクの評価

上のようにして得られた推定分布の VaR 等をリスク量とする場合、考えている値のランダム性に起因するプロセスリスク (process risk) は含まれているものの、パラメータの推定誤差に起因するパラメータリスク (parameter risk) やモデルの選択の誤りに起因するモデルリスク (model risk) は含まれていない。

この発表では、真の分布が正規分布または対数正規分布であると仮定した状況において、プロセスリスク、パラメータリスク、モデルリスクをすべて考慮した Bayes 推定によるリスク評価の具体的な方法を、実装例も含めて提示する。