

厚生労働科学研究費補助金（労働安全衛生総合研究事業）
（総括）研究報告書

経済情勢等が労働災害発生動向に及ぼす影響等に関する研究：
多変量時系列解析による数理モデルの開発と検証

代表研究者 松田文子（公財）大原記念労働科学研究所

研究要旨

マクロ経済学・金融工学等で応用されている多変量時系列解析手法(Kariya, 1993)を用いて、経済情勢が業種別労働災害の発生に及ぼす影響を明らかにすることが最終目的である。平成 28 年度は全体会合を 5 回（2016 年 10 月、12 月、2017 年 1 月、2 月、3 月）開催し、フォーカスグループ手法により、時系列モデルに投入する主要アウトカム・要因の選定を行った。また、解析に用いるデータ期間、および欠損値補完の処理方針（多重代入法、移動平均、変数変換処理など）について適宜ディスカッションを行い、主要な経済指標に関する変数の定義を行った。

A. 研究目的

労働災害(労災)は長期的には減少しているが、小売・飲食業や保健衛生業などの第三次産業では増加傾向にある。第 12 次労働災害防止計画においても、重点業種別の対策が提唱されているが、労働を取り巻く諸環境の要因（経済情勢、産業構造の変化、就業形態、自然・気象条件、産業技術革新等）が及ぼす影響について科学的根拠に基づく解析はほとんど行われておらず、行政政策評価に資する知見が切望されている。

そこで、本研究ではマクロ経済学・金融工学等で応用されている多変量時系列解析手法(Kariya, 1993)を用いて、経済情勢が業種別労働災害の発生に及ぼす影響を明らかにすることが最終目的である。

本研究は、5 つの研究班によって検討された内容を相互に活用しながら進める形態をとっている。

数理モデル班では、主に時系列モデルに投入する主要アウトカム・要因の定義の設定と生成、解析プロトコル手順の確立と数理モデルの構築、そしてモデル適合度評価の実施を担っており、平成 28 年度は、経済情勢班と協力連携し、主要アウトカム・要因の選定と定義の設定を行う。

労働経済班では、労働条件、業務内容、分業のあり方といった労務管理の変化に影響する指標の検討を行い、あわせて労働行政における労務管理の変化を認識・規制するプロセスについても検討する。以上をふまえ、多変量解析の指標選定およびトライアル解析を行い、課題を整理する。

労災分析班では、主に主要アウトカムとなる労災指標について、利用可能な変数の検討を行う。また、労災指標の時系列データを作成する際の問題点と課題について整理する。

気象天災班では、気象や天災に関係する指標の収集を行い、利用可能な変数の検討を行うとともに、課題を整理する。

B. 方法

全体会合を 5 回（2016 年 10 月、12 月、2017 年 1 月、2 月、3 月）開催した。各研究班で調査した各種指標を持ち寄り、各指標の利用可能性についてプレーストーミングを行った。各指標はデータ期間、データ密度（年単位・四半期単位・月単位など）、データの質（発行元や信頼性）、データの利用可能性（入手先）、データ加工の手間、データ欠損の度合いの 6 側面で検証を行い、最終的に投入する変数の定義方法および優先度について議論を重ねた。

なお本研究では、国が提供・公開している各種データ資源、統計法等、法令の規定に基づく調査データ（連結不可能匿名化後の統計データ）などオープンデータを主に扱うため、文部科学省・厚生労働省「人を対象とする医学系研究に関する倫理指針（平成 26 年 12 月 22 日）」は適用外である（個々の研究対象者からデータを収集することは行わない）。

C. 研究結果

1. 変数の抽出

最終的に抽出された変数は計 212 変数であった。それらオープンデータの利用可能期間は 1947 年（昭和 22 年）～2017 年（平成 29 年）の約 70 年の期間にまたがっているが、各変数により収集期間が異なるため、最大公約数的な期間を検討した結果、多くの経済系の指標は 1973 年以前は利用不可能であることから直近 50 年間程度が年単位デー

タの利用可能な最大の範囲であることが分かった。また、天災・自然災害や象徴的な経済事象（リーマンショックなど）など、経済動向の迅速な影響が反映される可能性も考えられるため、月単位の粒度を細かくした時系列変動解析についても行えるよう、月単位のデータセットの生成可能性についても検証を行った。月単位で利用可能なデータ利用可能期間は直近 25 年間程度と考えられた。以上より、暫定的に年単位データは 1973～2012 年前後の 50 年間、月単位データは 1992～2012 年の約 20 年間を対象データ期間とし、データセットの整備を進めることで合意した。

また、各データについては利用不可能な年または期間が存在する変数もある。年単位データの解析可能期間と想定される 1973～2012 年のデータセットにおいて、完全データの変数（当該期間に欠損がない変数）は全変数の 27%に過ぎなかった。また、同じ統計であっても出処が異なると当該年の統計値が異なるものも存在し、欠損値補完の処理方針（多重代入法、移動平均、変数変換処理など）についても研究組織メンバー内で適宜ディスカッションを行った。時系列データで何年未満の欠損をどのような統計手法で補完するのかについては更なる議論が必要であった。現段階では同一年でも出処により値が異なる変数は、変数 XXXXa, XXXXb のようにコード化しデータシート内では併記しておくこととした（またデータシート内で該当箇所を分かるようにマークしておく）。

抽出した変数については、経済情勢関連については、主に国内総生産、産業別労働生産性、景気動向指数、設備投資・技術革新指数（機械受注指数など）、各種経済指数（企業物価指数、失業率、鉱工業生産指数、消費者物価指数など）など。労働経済関連については、産業構造、就業形態（非正規雇用、外国人労働、高齢労働等）や社会政策、労働力人口（業種別就労人口）などを示す各種指標を、気象・天災関連については、自然災害（台風、豪雨など発生日時および被害数）、気温のほか、大震災など経済活動にも影響を与えられられる事象（震度 5 弱以上、死傷者 100 人以上の災害）については月単位データとして抽出・データ生成を行った。また、従属変数（アウトカム）として利用する労働災害関連の指標については、業種別度数率、強度率、労働損失日数、労災申請件数などの利用を想定しているが、代理アウトカム指標の利用可能性も視野に人口動態統計から死因別死亡率などの情報も収集を行った。

2. 経済動向を示す主要変数

本研究課題として扱う経済動向としては、いわゆる一般的に扱われる事の多い Index を主要変数として位置づけ、その他補助的変数の選定をモデル投入時に探索的に行うこととした。

D. 考察

暫定的に年単位データは 1973～2012 年前後の 50 年間、月単位データは 1992～2012 年の約 20 年間を対象データ期間とし、データセットの整備を進めることで合意した。

また、5 回の全体会合を重ねる中で、個々の指標について質的側面についても適切に理解をしておく必要性が提案され、議論を重ねてきた。例えば、労働災害統計自体、1 次・2 次産業の減少と 3 次産業の増大など産業構造の変化に伴い 70 年代以降は働き方が変革していること、また各種法制度の整備の状況や労働災害の質的变化（メンタル疾患の増加など）などもあり、単に量的側面だけで捉えることが困難になりつつある。各指標について、それぞれ想定される背景因子が存在すると考えられるため、モデル投入を行う前に適切な調整法（モデル人口調整、死亡率調整、基準年調整などによる各種補正值）を検証することも必要となるため、平成 29 年度では各指標について下記プロファイル情報を一覧整理することとした。

< 各指標のプロファイル情報 >

- ・ 出処情報
- ・ サンプリング方法：データ収集方法、分母・分子の定義ほか。
- ・ 指標の求め方・算出方法
- ・ 基準年の有無
- ・ 変数の特徴：トレンド、定常・非定常、ランダムウォーク、周期性に関する傾向
- ・ 変数の調整方法（既存研究での事例検索）
- ・ 要配慮点・特筆すべき事項：
- ・ 時系列変動データ（図示）

これら各指標の特性を把握し、モデル投入する変数の優先度を付け、基礎解析として時系列解析へ変数投入すべきかの検証を行うこととした（ホワイトノイズ解析、交差相関、自己相関分析）。また変数の特徴としては、定常・非定常時系列、トレンド・ランダムウォーク性の解析、スペクトル解析による時系列推定モデルの検討（Holt トレンド、Brown トレンド、減衰トレンドなど）を行い、各変数の持つ時系列情報の特性を事前に明らかにしておくこととした。

現在、200 以上の指標（変数）を収集しており、これらの指標を多様に組み合わせることで、モデルの精度を上げることを目指しているが、その前段階として、現状のローデータをそのまま投入するのではなく、各指標のプロファイルを整理し、適切な補正を行うデータの下処理が重要といえよう。なお、現在収集済みの指数を見直していく中で、解析の枠組みもアプローチを分ける必要性が

明らかになってきた。すなわち、本研究課題で扱うデータはマクロデータによる解析のため、利用可能な指標が一定期間存在しない変数については扱うことができない。そこで、解析の粒度設計として長期的なトレンドで解析するアプローチと、即時的反応が期待される変動指数を用いた短期的アプローチの2つに分けて、経済動向が労働災害に与える影響を調べる必要があると示唆された。長年にわたって継続的に蓄積された指標については年単位、時代の要請により指標の重要性が提唱され、近年、取得され始めた指標で月単位で利用可能なものもあり、これら2つの異なる粒度でアプローチすることで、説明力を高めて行くことを目指している。

E. 結論

経済情勢が業種別労働災害の発生に及ぼす影響を明らかにするために、主に時系列モデルに投入する主要アウトカム・要因の抽出と定義の設定を行った。現段階で抽出された変数は計200以上におよび、それらオープンデータの利用可能期間は1947年(昭和22年)~2017年(平成29年)の約70年の期間にまたがっていた。各変数により収集期間が異なるため、最大公約数的な期間を検討した結果、暫定的に年単位データは1973~2012年前後の50年間、月単位データは1992~2012年の約20年間を対象データ期間とするのが合理的と判断された。ただし、完全データの変数(当該機

関に欠損がない変数)は27%に過ぎないこと、また自己回帰モデルによるトライアル解析からも死亡災害件数の予測には自己回帰モデルは不十分なことから、200以上の指標(変数)について各プロファイルを明らかにして、モデル投入を行う前に適切な調整法を検証することが重要となることが示唆された。

F. 健康危険情報

該当なし

G. 研究発表

1. 論文発表

平成28年度はなし

2. 学会発表

平成28年度はなし

H. 知的財産権の出願・登録状況

1. 特許取得

該当なし

2. 実用新案登録

該当なし

3. その他

該当なし