

若手研究者の会企画集会 「農業気象分野研究における統計解析 －実験計画法の理論と実践－」開催報告

臼井靖浩*・伊川浩樹**・大久保さゆり***

(* 農研機構 北海道農業研究センター
** 農研機構 農業環境変動研究センター
*** 農研機構 東北農業研究センター)

1. 概要

日時：2017年3月28日(火) 17:30～19:00

場所：北里大学十和田キャンパス本館B棟2階 (B会場)

主催：若手研究者の会

趣旨説明 臼井靖浩

(若手研究者の会代表幹事：

農研機構 北海道農業研究センター)

講演：

1. 「実験計画法に則った研究と観察研究の違い」
大東健太郎 (農研機構 農業環境変動研究センター)
2. 「対話形式によるデータ解析や表現方法についての議論」
長谷川利拡 (農研機構 東北農業研究センター)

2. 企画集会の趣旨

日本農業気象学会には、気象要素をはじめとする環境計測や気候変動が及ぼす将来予測など、時系列データ解析や確率統計モデルを用いた研究を端緒とする若手研究者が多い。一方で、多くの学会員が行う研究において、土壌、大気、作物といった境界領域を取り扱う際、作物の環境応答を実験的に検証する能力も必要とされる。そのため、いざ自ら試験区を設計し、実験結果を解析する際、統計解析でつまづいてはじめて、統計の専門家に駆け込む事例が多いと聞く。そこで、本集会では、実験計画を立てる前の注意点を統計の専門家にお話いただき、対話形式によって、データ解析や表現方法について議論した。

3. 講演内容

3.1 「実験計画法に則った研究と観察研究の違い」

大東氏からは実験計画を立てる際の注意点と観察的研究から得られる結果に対する、統計的解析手法と留意点を紹介していただいた。講演は、実験を始める前に統計解析を含めた実験計画を立てる必要性があることの重要性を強調されることから始まった。統計解析を行うにあたっての留意点を Fisher の三原則 (無作為化・反復・局所管理) に基

づいて説明した。また観察的研究には様々なバイアスが潜むことと、Confounding effect (交絡する影響) を取り除く必要性と指針を示した。

まず、実験後に統計処理について考えることは取り返しのつかない失敗をした後に原因を探るようなものであるという Fisher の考えを紹介し、サンプルサイズ (反復) の算定と無作為化のための工夫を怠ってはならないことを指摘した。ある実験に対してどの程度のサンプルサイズが必要であるかは、おそらく研究者の誰もが一度は疑問に思うことではないだろうか。サンプル数は、試行的データの分散、目標にする有意水準に加えて科学的に意味のある差が決まれば自動的に算出される。リソースが限られる場合においては分割法などを取り入れて実験系を工夫することを推奨した。

次に観察的研究における問題点と因果関係を導く手法について紹介した。観察的研究には、特定のグループが非協力的になる未回答者バイアス、サンプルが死滅することなどによるネイマンバイアス、誇張した事例をサンプルするバークソンバイアスなどのバイアスがあることを気に留めておく必要がある。現在では、観察的研究から因果関係を調べるためには構造的アプローチと統計的モデルアプローチが用いられている。因果構造を描写する構造的モデルの実用性は高く、力学ダイアグラムを描き、交絡的影響を調べることで、ターゲットとする因果関係の統計的有意性を検討することが重要であると述べた。

以上のように、Fisher の三原則に見られる統計解析における基本的理念から始まり、観察的研究における因果関係を統計から推定する最先端の統計学的手法を紹介された。最後に演者は、10年後には因果関係が統計的に有意であるかを検出する傾向スコア (IPW) などの概念が、観察的研究によって論文化を進める上でも必要とされるのではないかと述べられた。

3.2 「対話形式によるデータ解析や表現方法についての議論」

長谷川氏は、統計解析から導かれる有意性に対する解釈と実験単位のことについて説明され、実験単位に関しては、参加者を含めて、チャンバーの数が限られる場合や高度な気象観測など測定が一か所に限られる場合などの議論へと発展した。

まず、有意性の基準として一般的に使われる P 値とは何かから始まり、 P 値による統計的有意性の判断について説明された。 P 値が 0.05 をやや上回るが 0.05 に近い場合、結果に対してどのような解釈をするべきだろうか。Wasserstein & Lazar (2016) は次のように述べている。“Scientific conclusions and business or policy decisions should not be based only on whether a p -value passes a specific threshold.”つまり P 値がある基準より高い場合でも、論理的に考えて正しい選択をすることの重要性について説明した。

次に、実験単位に関して、階層構造を持つ実験系とその反復を中心に説明がなされた。実験系が階層構造を持つ場合、完全にランダムサンプルとした統計解析を適用してはいけないと述べた。例えば、灌漑のような比較的コントロールが難しい実験においては、一つの灌漑試験区の中に、複数のサンプリングを設けることが一般的であろう。このような場合、それぞれのサンプルは同一の灌漑試験区から得られたデータなので、灌漑の効果を完全にランダムとしてはならず灌漑の効果自体は一つの灌漑試験区の中のサンプル数に依存しないことを言及した。実験の反復の必要性に関してはある程度、それぞれの研究分野における一般的見解にも左右されることについて、参加者を交えて議論した。

以上のように、統計解析における結果の解釈は必ずしも一様にされるべきものではなく、論理的な判断を要することと、実験単位に気を付けて実験計画を立てることの重要性についての議論が交わされた。

4. 所感

本講演に関して、大東氏には統計解析を見越した実験計画の立て方を、長谷川氏には実験単位に気を付けた実験計画と統計解析結果の解釈についてお話を頂けた。両演者からの話題に共通する最たる事項は、実験計画を入念に立てることの必要性であるように思う。大東氏が述べられたように、実験が終わった後に解釈に都合のよい統計手法を探す、いわゆるショットガンの手法にならないような注意も必要であろう。

今回開催した企画集会には、約 80 名もの参加者があった。企画に際しては、統計で困った際に、「誰に、どんな相談をすればよいのか？」という点を解決するきっかけを作りたいという動機があった。企画集会の後、さっそく相談している参加者もあり、目的の一つは達成したといえる。今回の盛況を機に、また今後も統計に関する講座を開いてほしいとの要望も寄せられているため、統計に関する研究会の立ち上げなど、継続して開催できる仕組みを模索したいと考えている。

引用文献

Wasserstein RL, Lazar NA, 2016: The ASA's Statement on p -Values: Context, Process, and Purpose. *The American Statistician* **70**, 129–133.