

生態学・進化生物学のメタ解析のための統計モデル

中川 震一¹・久保 拓弥²

(受付 2015 年 6 月 30 日；改訂 2016 年 3 月 2 日；採択 3 月 25 日)

要 旨

生態学・進化生物学の分野において、いまやメタ解析は多くの一次研究(primary study)を定量的に統合するもっとも有望な手法となっている。この手法はもともとは医学・社会科学の分野で発展してきたもので、それは固定効果(fixed effects)モデルやランダム効果(random effects)モデルなどの適用から始まった。メタ解析で扱うデータとは効果量の集まりであるが、生態学・進化生物学の分野ではこれらはより不均質(heterogeneous)かつ相互依存的(inter-dependent)であるという特徴を持つので、効果量間の独立を仮定している上にあげた従来のメタ解析モデルでは、うまくあつかえない。生態学・進化生物学分野におけるメタ解析では、一次研究内での効果量の非独立性、あるいは対象となる生物種(species)間の系統学的な近縁性といった非独立性(相関構造)をあつかわなければならないことが多い。これらの非独立性を扱うために提案されたメタ解析の統計モデルを紹介する。系統学的な比較法をくみこんだマルチレベルモデル、すなわち系統学的マルチレベルメタ解析は生態学・進化生物学分野で頻出するデータを解析するのに適している。またメタ解析の不均質性 I^2 とメタ回帰の R^2 の概念についても検討する。メタ解析のモデルは発展しつつあるが、生態学・進化生物学分野ではその利用は進んでいない。この分野の研究者たちに対する実効性のある教育プログラムが必要である。

キーワード：システムティックレビュー、定量的研究、データ統合、階層モデル、混合効果モデル、系統樹。

1. はじめに

Glass が「メタ解析」なる用語を作りだして、この統計的な技術の現代的な用法を定式化したのは、40 年ほど前のことである(Glass, 1976, 2015)。最初に社会科学・医学の分野で、メタ解析は複数の一次研究(primary study)を偏りなく統合するために活用された(たとえば Egger et al., 2001; Cooper et al., 2009)。生態学・進化生物学の分野では、1990 年代の始めごろからメタ解析の手法が使われるようになった。この分野におけるメタ解析の有望性を予見していた、影響力ある総説 Arnqvist and Wooster (1995)によれば、最初のメタ解析がなされたのは 1991 年ということになる。近年では、生態学・進化生物学におけるメタ解析の書籍が出版されている(たとえば Koricheva et al., 2013; Nakagawa and Poulin, 2012)。とくに“*Handbook of Meta-analysis in Ecology and Evolution*” (Koricheva et al., 2013)は、この分野におけるメタ解析の定着と成熟を示すものといえるだろう。

¹ Evolution & Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia

² 北海道大学大学院 地球環境研究科：〒060-0810 札幌市北区北 10 条西 5 丁目

そもそも医学・社会科学の分野で発展してきたメタ解析の方法が、生態学・進化生物学のデータ解析には必ずしも適したものではないとしても不思議ではない。この分野に独特な問題の数々が指摘されてきた (Arnqvist and Wooster, 1995; Gurevitch and Hedges, 1999; Nakagawa and Santos, 2012)。医学・社会科学分野での研究とのもっとも顕著なちがいは、これらの分野では「ヒト」という一種類の生物だけをあつかっているのに対して、生態学・進化生物学では幅ひろい生物種・系統を研究対象としているところである。このため、多くの場合、生態学・進化生物学におけるメタ解析を実施するために、不均質さの大きい一次研究をたばねる方法が必要となる。その結果として、この分野では医学・社会科学とは異なる統計モデルが必要とされている。

ここで生態学・進化生物学におけるメタ解析の一例を簡単に紹介してみよう。Cleasby and Nakagawa (2012) は、スズメ目の社会的一夫一妻制の鳥類におけるつがい内・外の父性と年齢の関係について調べている。このような解析によって、オスの齢と「浮気」の関係が明らかになれば、鳥類における社会構造と性淘汰の進化の理解に寄与できるからである。このメタ解析は、キーワード検索で見つけた 1210 篇の解析候補となる一次研究のうち、メタ解析に利用可能な 61 篇を対象とするものであり、スズメ目 35 種・40 集団が含まれている。これらの全 35 種は同じ系統樹上で系統的距離の長短を評価することができ、つまり比較対象となる 35 種のうち 2 種を選んだ場合に、系統的に近い・遠いといった要因を考慮しなくてはならない。また鳥の種ごとに異なるタイプのデータがとられていて、その組み合わせが種ごとに異なるのが特徴である。これらの各一次研究で示された結果のうち、三種類の効果量 (オスの齢とつがい内父性の関係、オスの齢とつがい外父性の関係、そして齢による父性の変化) をひとつずつ、それぞれメタ解析している。ここでいう父性とは、巣の中の子供がどのオスの遺伝子を持つかに着目した観測値であり、たとえばつがい内父性であれば、巣の社会的親とは異なる遺伝子をもつ子供の割合 (あるいはそういう子供がいるかないか) が使われる。つがい外父性は自分の巣の外で育てられている子供の個数である。また、「齢による父性の変化」とはつがい内交尾しているオスとつがい外交尾をしているオスの齢の差である。このような多数の一次研究で得られた効果量とその分散をまとめて説明できるような統計モデルを構築し、「メタ」な効果量を推定する手法がメタ解析である。

この総説では、生態学・進化生物学で使われているメタ解析の統計モデルの概要を紹介する。最初に、メタ解析の基本となる 2 つのモデルである、固定効果モデルとランダム効果モデルを説明し、その問題点を指摘する。つぎに、生態学・進化生物学の分野で発展してきた、より複雑ではあるが研究対象に適合した統計モデルを紹介する。不均質性の解析とメタ回帰の諸問題について検討し、実装・論文公表バイアス・データの欠測・多変量解析などについてもふれてみたい。

2. 従来から使われていたメタ解析モデル

メタ解析とは、興味の対象となる変数が効果量 (effect size) の推定値であり、それぞれの効果量に対応する標本分散の逆数を加重値とする、加重平均の特殊なものである。生態学・進化生物学では、次にあげる 4 種類の効果量がよく使われている (Nakagawa and Santos, 2012; Koricheva and Gurevitch, 2014) :

- (1) 応答の比率：対数をとって $\ln RR$ と書かれる
- (2) 標準化された平均の差：Cohen の d や Hedges の d
- (3) 相関係数：これは Fisher の z 変換をほどこした Zr
- (4) オッズ比：対数をとって $\ln OR$ と書かれる

これらのよく使われる効果量では、一次研究 j ごとに異なる標本分散 σ_j^2 も得られる。たとえば、 n を標本数とすると、 Z_r の標本分散は $1/(n-3)$ である (Nakagawa and Cuthill, 2007; Borenstein, 2009)。メタ解析では固定効果モデルとランダム効果モデルがよく使われているが (Hedges and Olkin, 1985; Hedges and Vevea, 1998)、まずはこれらを生態学・進化生物学の研究に応用してみる、という方向性で話を進めてみよう。

2.1 固定効果モデル

固定効果モデルを数式で書くと以下ようになる：

$$(2.1) \quad \begin{aligned} z_j &= \mu + m_j, \\ m_j &\sim N(0, \sigma_j^2), \end{aligned}$$

ここで z_j は第 j 研究における効果量の値 ($j = 1, \dots, N_{\text{study}}$; N_{study} は研究の個数)、 μ はメタ解析の全平均である。 m_j は一次研究 j におけるサンプリング誤差で、平均ゼロであり、一次研究 j ごとに異なる分散 σ_j^2 となる正規分布にしたがう。前に説明したように σ_j^2 は既知である。このモデルでは、解析対象となる全ての一次研究の真のメタ解析平均が存在すると仮定している。

生態学・進化生物学のメタ解析では、この仮定はほとんどいつもみだされていらない。その理由は、一次研究の実験計画・個体群・生物種・系統などのばらつきが大きすぎるためである。この固定効果モデルはメタ解析の普及初期にはよく使われていたが、その後は使われなくなりつつある (Nakagawa and Poulin, 2012; Mengersen et al., 2013)。医学・社会科学でも同じように減少傾向になっているようだ (Higgins et al., 2009)。しかしながら、メタ解析でとりあげる一次研究の個数が少ないときには、固定効果モデルを使うほうがよいだろう。一次研究の個数 N_{study} が小さいときには、次に説明するランダム効果モデルを使った分散の推定がうまくいかず、その悪影響は他のパラメーターの推定にもおよんでしまうからである (Mengersen et al., 2013)。

2.2 ランダム効果モデル

ランダム効果モデルは次のように書ける：

$$(2.2) \quad \begin{aligned} z_j &= \mu + u_j + m_j, \\ u_j &\sim N(0, \sigma_u^2), \\ m_j &\sim N(0, \sigma_j^2), \end{aligned}$$

式中の z_j は第 j 研究における効果量の値、 μ はメタ解析の全平均、 u_j は一次研究 j に固有な効果であり、これは平均ゼロで一次研究間の分散 σ_u^2 固定モデルの場合と同じく、 m_j は一次研究 j におけるサンプリング誤差であり、一次研究 j ごとに異なる分散 σ_j^2 (これも固定モデルと同じで既知) の正規分布にしたがう。他の記号は固定効果モデルと同じである。この σ_u^2 を推定する方法があり、例えば Hunter-Schmidt 推定量、Hedges 推定量、DerSimonian-Laird 推定量、REML 推定量などであり、ここでは説明しないが、興味のある読者は Sánchez-Meca and Marín-Martínez (2008) などを調べてほしい。

生態学・進化生物学のメタ解析では、このランダム効果モデルが選ばれることが多く (Mengersen et al., 2013)、後述するいろいろな限界があるにもかかわらず、もっともよく使われている。その理由は、メタ解析のランダム効果モデルは、一次研究たちの全般的なトレンド (overall trend, メタ解析全平均) だけでなく、一次研究間の不一致性なども推定可能であり、これらは不均質性 (heterogeneity, 次項で定義する) とよばれている。生態学・進化生物学のメタ解析では、ひとつの研究の中でいろいろな観点から調べるので、全平均だけでなく不均質性の大きさを見積も

ることが必要となる。全般的なトレンドは、この不均質性の大きさに左右されている。

2.3 不均質性の解析

これまで、一次研究が不均質(σ_u^2 が非ゼロ)であるかどうかを調べるために、Cochran (1954)が定義した Q 検定が使われていた。しかしながら、この Q 検定は統計学的な有意性を示すだけで、少なくとも直接的には不均質性の大きさを数値的に示せない。そこで、のちにこの難点を解決するために、0から1までの値をとる I^2 統計量が提案された(Higgins and Thompson, 2002)。この I^2 は以下のように定義される：

$$(2.3) \quad I^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_m^2},$$

$$\sigma_m^2 = \frac{(k-1) \sum \sigma_j^{-1}}{(\sum \sigma_j^{-1})^2 + \sum \sigma_j^{-2}},$$

ここで σ_u^2 は一次研究間の分散、 σ_m^2 は「典型的な」一次研究内の分散であり、研究固有のサンプリング誤差 σ_j^2 を組み合わせて得られる。式(2.3)はHiggins and Thompson (2002)で提案され、 σ_m^2 の定義としてはもっともよく使われているものだが、これとは別の推定量もある(たとえばCheung, 2014)。式(2.3)を見ると、 $(\sigma_u^2 + \sigma_m^2)$ が応答変数(効果量 z_j)の全分散なので、 I^2 は効果の大きさが一次研究ごとに異なることによる、分散の割合であることがわかる。つまり、 I^2 はサンプリング誤差によらない変異だと言える。Higgins et al. (2003)では I^2 推定値の0.25, 0.50, 0.75の区切りが、それぞれ小さい・普通・大きい不均質性に対応すると考えればよいとしている。このような I^2 の基準値はCohenが提案した相関係数と標準化平均の基準値を参考にしたものであり、それは r 値が0.1, 0.3, 0.5そして d 値が0.3, 0.5, 0.8の区切りで、効果が小さい・中くらい・大きいと対応させている(Cohen, 1988)。

この不均質性指数 I^2 を使った解析は、固定効果モデルからの大きな改善である。しかし、このランダム効果モデルには大きな欠点がある。それは、それぞれの一次研究の中の複数の効果量のあいだの非独立性をあつかえない、という点である。このような非独立性、あるいは効果量に関する他の非独立性は、生態学・進化生物学のメタ解析ではよく見られるものである。そこで、ランダム効果モデルよりさらに新しい統計モデルとして、効果量間の非独立性をうまくあつかえるような、マルチレベル(あるいは階層)混合効果モデル(Raudenbush and Bryk, 2002; Gelman and Hill, 2006)が使われるようになってきた。

3. マルチレベルメタ解析

生態学・進化生物学の研究者たちが、マルチレベルメタ解析をするようになったのはごく最近である(ただしLiermann and Hilborn, 1997)。それまでは、以下のような方法で、独立ではない問題に対処もしくは「ないこと」にしていた：

- (1) 独立を仮定する
- (2) それぞれの一次研究内で平均化した対応のある効果量(dependent effect size)を用いる
- (3) ひとつの効果量だけに注目する(たとえばCheung, 2014)

ひとつめは明らかにまちがいであり、第一種の過誤の確率を増大させる。二番目・三番目は必ずしもまちがいではないが、データから言えることが少なくなり検定力も大きく減少する。これにたいして、メタ解析でマルチレベルモデルを使うと、すべての問題をうまく回避できる。生態学・進化生物学のメタ解析で頻繁に扱う非独立性(依存性)に対処できる。これには二種類あ

り、ひとつめはすでに述べたように、それぞれの一次研究内における効果量の非独立性であり、もうひとつは系統学的な近縁関係がもたらす非独立性である。前者は他の分野のメタ解析でもとりあつかわれるかもしれないが、後者は生態学・進化生物学のメタ解析だけに見られるものである。

3.1 ひとつの一次研究から得られる複数の効果量

一次研究内で複数の独立ではない効果量がある場合、ランダム効果モデル(式(2.2))を拡張することで簡単なマルチレベルモデルとなる：

$$(3.1) \quad \begin{aligned} z_i &= \mu + u_{j[i]} + m_i, \\ u_j &\sim N(0, \sigma_u^2), \\ m_i &\sim N(0, \sigma_i^2), \end{aligned}$$

ここで z_i は第 j 番目の一次研究における第 i 効果、 μ はメタ解析の全平均、 u_j は第 j 番目の一次研究 ($j = 1, \dots, N_{\text{study}}$) における第 i 効果 ($i = 1, \dots, N_{\text{effect}}$)、すでに説明した固定またはランダム効果モデルの場合と同様に、 m_i は一次研究 i におけるサンプリング誤差であり、平均ゼロかつ既知の標準偏差 σ_i の正規分布にしたがう。 N_{effect} は効果量の個数であり N_{study} より大きくなくてはいけない。他の記号はランダム効果モデルと同じである。式(3.1)の記法について補足説明をすると、 z_i の式の右辺に $u_{j[i]}$ のような添字 $j[i]$ がついた確率変数がある場合、 z_i が第 j 番目の一次研究における第 i 効果であるとする。また u_j が平均ゼロ・標準偏差 σ_u にしたがうとする場合には、左辺 u_j の添字では $[j]$ が省略される。このような統計モデルの記法は Gelman and Hill (2006) によって提案され、近年では社会科学だけでなく、多くの学術分野におけるメタ解析モデルの記述に広く用いられているので、本稿でもそれにしている。

上で紹介したモデルは、それぞれの一次研究内で独立ではない効果量をうまくあつかえるものなのだが (Evans et al., 2010; Kamiya et al., 2014)、かならずしも正しいとは言えない仮定がなされている。その仮定とは、一次研究内の分散はサンプリング誤差の分散の一部だとしていることである(式(2.3))。一次研究内の分散とサンプリング誤差の分散を分離したモデルは、このように書ける：

$$(3.2) \quad \begin{aligned} z_i &= \mu + u_{j[i]} + e_i + m_j, \\ u_j &\sim N(0, \sigma_u^2), \\ e_i &\sim N(0, \sigma_e^2), \\ m_i &\sim N(0, \sigma_i^2), \end{aligned}$$

ここで e_i は第 i 効果量に固有な(それぞれの一次研究内の)ばらつきであり、それぞれ平均ゼロで一次研究ごとに異なる分散をもつ正規分布にしたがい、直線回帰モデルにおける残差項に相当する。他はすぐ上で説明した式(3.1)の統計モデルと同じである。

ひとつ注意すべきなのは、式(3.1)のマルチレベルモデルはこのモデル(式(3.2))よりも便利であることで、その理由は不均質性 I^2 が式(2.3)で推定できるからであり、このモデルで不均質性を定量化する方法はあとで紹介する。また、式(3.1)のモデルは、一次研究それぞれの分散を正確に推定できない(たとえば N_{effect} と N_{study} がそれほどちがわない)ときに、より適した方法かもしれない。

一次研究内の複数の効果量について、上とは別の頻出する問題がほかにもあり、それもまた無視されることが多い (Curtis and Queenborough, 2012)。その問題とは、2つのグループを比較する $\ln(RR)$ や Hedge の d といった統計量を使うときに発生する。Lajeunesse (2011) が指摘

したように、生態学あるいは進化生物学の実験において、無処理区をひとつ、そして処理区を複数設定するような研究で一般的な問題である。たとえば無処理区はひとつ、処理区は2つあるとしよう。すると2つの効果量が得られる。この2つの効果量は、どちらも無処理区からのずれなので、その非独立性は適切にモデル化されなければならない。この問題は、さきのマルチレベルモデル(式(3.1)と(3.2))ではうまくあつかえない、という点に注意してほしい。そこで、式(2.1)に相当する分散共分散行列を使ったモデルが必要になる：

$$(3.3) \quad \mathbf{m} \sim N(\mathbf{0}, \mathbf{M}),$$

これを式(3.2)のモデルで使うとすると、 \mathbf{m} は $\{e_i\}$ ($i = 1, \dots, N_{\text{effect}}$) のベクトルで、これは平均がゼロ行列 $\mathbf{0}$ で与えられ、 $N_{\text{effect}} \times N_{\text{effect}}$ の分散共分散行列 \mathbf{M} をもつ多変量正規分布にしたがうことになる。

ひとつの例として、 $N_{\text{effect}} = 3$ であるもっとも単純な \mathbf{M} を考えてみよう。ここでは3つのうち2つの効果量が無処理区と対比されている。そのような \mathbf{M} は次のように書ける：

$$(3.4) \quad \mathbf{M} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & 0 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix},$$

上の $\sigma_1^2, \sigma_2^2, \sigma_3^2$ はそれぞれの標本分散、そして $\rho\sigma_1\sigma_2 (= \rho\sigma_2\sigma_1)$ は第一と第二の効果量の共分散で、これらは無処理区の効果量に依存している (ρ は相関係数)。先にあげたよく使われている効果量の統計量についての $\rho\sigma_1\sigma_2$ の推定量は簡単なかたちで得られる ($\ln(RR)$ については Lajeunesse, 2011, Hedge の d については Gleser and Olkin, 2009)。たとえば $\ln(RR)$ の場合、 $\sigma_1^2, \sigma_2^2, \rho\sigma_1\sigma_2$ の推定量は次のように書ける：

$$(3.5) \quad \begin{aligned} \hat{\sigma}_1^2(\ln(RR)) &= \frac{s_C^2}{n_C \bar{x}_C^2} + \frac{s_{T_1}^2}{n_{T_1} \bar{x}_{T_1}^2}, \\ \hat{\sigma}_2^2(\ln(RR)) &= \frac{s_C^2}{n_C \bar{x}_C^2} + \frac{s_{T_2}^2}{n_{T_2} \bar{x}_{T_2}^2}, \\ \hat{\rho}\hat{\sigma}_1\hat{\sigma}_2(\ln(RR)) &= \frac{s_C^2}{n_C \bar{x}_C^2}, \end{aligned}$$

ここで s_C, s_{T_1}, s_{T_2} は、それぞれ無処理区・第1処理区・第2処理区の標本標準偏差であり、 n と \bar{x} はサンプルサイズと平均である。上の第3の式で共分散 ($\hat{\rho}\hat{\sigma}_1\hat{\sigma}_2$) つまり $s_C^2/n_C \bar{x}_C^2$ が最初の2つの式のどちらにも入っていて、第1処理区・第2処理区どちらもこれを共有していることがわかるだろう。

さて、ここで式(3.2)のマルチレベルモデルにもどって、ひとつあるいは複数の層を入れてみたい。たとえばメタ解析のデータセットには、複数の生物種 (species) が含まれているとしよう。式(3.2)はこのように書き直せる：

$$(3.6) \quad \begin{aligned} z_i &= \mu + v_{k[i]} + u_{j[i]} + e_i + m_i, \\ v_k &\sim N(0, \sigma_v^2), \\ u_j &\sim N(0, \sigma_u^2), \\ e_i &\sim N(0, \sigma_e^2), \\ m_i &\sim N(0, \sigma_i^2), \end{aligned}$$

ここで z_j は第 j 研究における効果量の値、 μ はメタ解析の全平均、 $v_{k[i]}$ は第 i 番目の効果量

の推定値に与える種 k 固有の効果 ($k = 1, \dots, N_{\text{species}}$; $N_{\text{effect}} > N_{\text{study}} > N_{\text{species}}$ であることに注意), v_k は平均ゼロで種固有の分散 σ_v^2 の正規分布にしたがうとする. また e_i は第 i 効果量に固有な(それぞれの一次研究内の)ばらつき, u_j は一次研究 j に固有な効果, m_i はサンプリング誤差である. 追加する層(stratum)は「生物種」に限定されるものではなく, 個体群や系統でもよい. このモデルが役にたつ場合もあるのだが, 系統関係が原因となって生じる効果量間の相関は考慮していない. この点に関して, Chamberlain et al. (2012)の最近の研究において, 多くの生態学・進化生物学分野のメタ解析を再検討し, 系統的な情報がメタ解析の結果を変えうることを示している. この研究によってメタ解析における系統関係の重要性が明らかになった.

3.2 系統関係が原因となる非独立性

進化生物学の分野では, 系統間の近縁関係を明示的にモデル化する比較法を使って生物間の形質を比較してきた長い歴史がある (Harvey and Pagel, 1991; Garamszegi, 2014). とくに, 線形回帰を使った系統的種間比較の方法は, 種間の形質進化において中心的な役割を果たしてきた. 系統学的比較解析で開発されてきた手法にもとづいて, Adams (2008)はメタ解析の固定効果モデル(式(2.1))に系統学的近縁性をくみこむ方法を提案し, あらたに“系統学的メタ解析”という用語を作った. その後, Lajeunesse (2009)が系統学的メタ解析をランダム効果モデル(式(2.2))に拡張した. これら2つの系統学的メタ解析モデルは次のようにあらわせる:

$$(3.7) \quad \begin{aligned} z_k &= \mu + a_k + m_k, \\ z_k &= \mu + a_k + u_k + m_k, \end{aligned}$$

上の μ はメタ解析の全平均, a_k は第 k 種の系統学的な効果(上の2つのモデルでは $N_{\text{effect}} = N_{\text{study}} = N_{\text{species}}$ となっていることに注意). また, m_k と u_k はそれぞれ, 一次研究 k におけるサンプリング誤差と k に固有な効果である.

$$(3.8) \quad \begin{aligned} m_k &\sim N(0, \sigma_k^2), \\ u_k &\sim N(0, \sigma_u^2), \\ \mathbf{a} &\sim N(\mathbf{0}, \sigma_a^2 \mathbf{A}), \end{aligned}$$

\mathbf{a} は長さ N_{species} の a_k のベクトルで, これは平均ゼロで分散共分散行列 $\sigma_a^2 \mathbf{A}$ の多変量正規分布にしたがい, ここで σ_a^2 は系統分散(phylogenetic variance), \mathbf{A} は系統樹から得られた種間距離をあらわす $N_{\text{species}} \times N_{\text{species}}$ の相関行列であり, これについてはあとで説明する. 複数の生物種の系統樹は, 分子データにもとづいて作られることが多い. たとえば, 鳥類と哺乳類についてはほとんどの種を網羅する系統樹が利用可能である (Bininda-Emonds et al., 2007; Jetz et al., 2012). ここでは $N_{\text{species}} = 3$ の場合について考えてみると, 行列 \mathbf{A} はこのように書けるだろう:

$$(3.9) \quad \mathbf{A} = \begin{bmatrix} 1 & f(g_{1,2}) & f(g_{1,3}) \\ f(g_{2,1}) & 1 & f(g_{2,3}) \\ f(g_{3,1}) & f(g_{3,2}) & 1 \end{bmatrix},$$

関数 f はこのあとで説明する「距離」の関数であり, $g_{i,j}$ は種 i と j を選んだときに, 系統樹の根からこの二種の最も新しい共通祖先までの距離である. この $g_{i,j}$ が1に近いほど種 i と j の系統間距離は小さい ($0 \leq g_{i,j} \leq 1$). ここで注意してほしいのは, 系統的メタ解析に使う場合, 系統樹の根と各末端の距離を決める単位距離が必要であり, 系統間の距離は超距離 (ultrametric) とすべきだということである. そのようにしなければ, 系統樹にもとづいて相関行列を作ることができないからである.

進化のブラウン運動モデルを仮定すると、 $g_{i,j}$ の関数 f は恒等関数(identity)であり $f(g_{i,j}) = g_{i,j}$ となる。ブラウン運動モデルでは表現型が自由に変化するのに対して、形質の値が特定の値から離れないような淘汰もありえる。このような安定化淘汰を仮定している例のひとつとして Ornstein-Uhlenbeck モデルがある。このモデルを採用した場合の関数 f の一例は $f(g_{i,j}) = \exp(-\alpha(1-g_{i,j}))$ であり、ここで α は進化的な制約の強さであると考えることができる。これら以外の進化モデルもありうるのだが、ここでは紹介しない。興味ある読者は Garamszegi (2014) を参照してほしい。

上で述べたように、式(3.7)で提案された系統学的メタ解析モデルでは、多くの場合、生物種のレベルで効果量が平均化されて $N_{\text{effect}} = N_{\text{study}} = N_{\text{species}}$ と仮定される。これは先にあげた一次研究内の効果量の平均化についての議論と同じことで、利用可能な情報と検定力が減少する。その解決法のひとつとして、Hadfield and Nakagawa (2010) は系統学的なモデルとマルチレベルモデルを組み合わせる方法を提案した。系統的マルチレベルメタ解析 (phylogenetic multilevel meta-analyses) のモデルは以下のように書ける：

$$(3.10) \quad \begin{aligned} z_i &= \mu + a_{k[i]} + w_{k[i]} + u_{j[i]} + e_i + m_i, \\ w_k &\sim N(0, \sigma_w^2), \\ u_j &\sim N(0, \sigma_u^2), \\ e_i &\sim N(0, \sigma_e^2), \\ m_i &\sim N(0, \sigma_i^2), \end{aligned}$$

$a_{k[i]}$ は第 k 種の系統学的な効果、 w_k は系統学的な効果 (a_k) とは無関係な第 k 種の i 番目の効果、 u_j は第 j 番目の一次研究における第 i 効果である。平均ゼロで分散 σ_w^2 の正規分布にしたがう。この w_k の式(3.6)で使われている $v_{k[i]}$ を区別することは重要である。どちらも種固有な効果であるのだが、 $v_{k[i]}$ では第 k 種における系統学的・非系統学的な効果の両方をあらわしている。多くの生態学・進化生物学のメタ解析では、複雑かつより妥当な系統学的マルチレベルモデルを使っている(たとえば Cornwallis et al., 2010; Lagisz et al., 2013)。また、 e_i は第 i 効果量に固有な(それぞれの一次研究内の)ばらつき、 m_i は一次研究 i におけるサンプリング誤差であり、それぞれ式(3.2)と(3.1)の説明を参照してほしい。

理論的には、式(3.10)のモデルがおそらく種間メタ解析にもっとも適したモデルだろう。しかしながら、このモデルを意図的に使っていないメタ解析もあり(たとえば Weir et al., 2011)、そのかわりに式(3.1)–(3.6)で示したより単純なモデルを使っている。その理由のひとつは、行列 A を作れるような系統樹が得られないためだ。生態学・進化生物学のメタ解析ではまったく異なる生物種の集合、たとえば昆虫・魚類・哺乳類などを扱う場合があげられる。とくにあまり研究されていないいくつかの種では、分類学的な情報は得られても系統学的な情報が得られない場合があるかもしれない。Hadfield and Nakagawa (2010) ではこの問題にも対処する方法を考えていて、比較生物学(Harvey and Pagel, 1991)でよく使われている伝統的な分類学的モデルと、この系統学的メタ解析をくみあわせる方法を提案している。たとえば、科(family)より「下」のレベルでは系統学的な関係がわかっているとしよう。すると科より下の系統樹をくみこんだモデルは、すぐ上で説明した式(3.10)を少しだけ変えて、次のように書ける：

$$(3.11) \quad \begin{aligned} z_i &= \mu + a_{l[i]} + q_{l[i]} + w_{k[i]} + u_{j[i]} + e_i + m_i, \\ q_l &\sim N(0, \sigma_q^2), \end{aligned}$$

ここで q_l は平均ゼロで分散 σ_q^2 の正規分布にしたがい、 $q_{l[i]}$ は l 番目の科に固有な効果であり、第 i 番目の効果量に影響をおよぼしていて、系統学的な効果 a_l とは別のものである。科の番号は

$l = 1, \dots, N_{\text{family}}$ (N_{family} は分類学上の科の個数) であり, $N_{\text{effect}} > N_{\text{study}} > N_{\text{species}} > N_{\text{family}}$ となっていることに注意してほしい. 以前と同様に, w_k は種 k に固有な効果, u_j は一次研究 j に固有な効果, e_i は第 i 効果量に固有な(それぞれの一次研究内の)ばらつき m_i は一次研究 i におけるサンプリング誤差である.

ここまで, 複雑ではあるけれど妥当なモデルを紹介してきたが, 現実にはデータサイズが限られているといった理由で, より簡単なモデルを利用することになるかもしれない. 統計モデルが複雑になり, より多くのパラメーターがくみこまれると, これらを推定するためにより多くのデータが必要になる. たとえば, マルチレベルモデルに新しく層を追加する場合には, 高次でのサンプル数(たとえば N_{study})より低次のそれ(N_{effect} など)は多くなるようにしなくてはならない. このような制約があるので, 統計モデリングでは理論的には良いと考えられるものと, 使えるデータで実現可能なものの中でバランスをとらねばならない.

3.3 マルチレベルモデルの不均質性解析

先に I^2 が不均質性をどのように定量化しているかを説明してみた. ひとたびこの I^2 がレベル内相関(intra-class correlation, ICC)の一種なので, I^2 の概念をさらに拡張できる(Nakagawa and Schielzeth, 2010). Nakagawa and Santos (2012)では I^2 はそれぞれのレベルで, サンプリングエラー m と分離して推定する方法を提案している(同様に Cheung, 2014)たとえば, 式(3.10)のモデルでは全分散 σ_t^2 は以下のように分割できる:

$$(3.12) \quad \sigma_t^2 = \sigma_a^2 + \sigma_w^2 + \sigma_u^2 + \sigma_e^2 + \sigma_m^2,$$

これらの記号を使って, I^2 を各レベルで推定できる. 以下に列举すると, 系統学的レベルでは $I_a^2 = \sigma_a^2/\sigma_t^2$, 生物種レベルでは $I_w^2 = \sigma_w^2/\sigma_t^2$, 研究間レベルでは $I_u^2 = \sigma_u^2/\sigma_t^2$, 研究内レベルでは $I_e^2 = \sigma_e^2/\sigma_t^2$ となる. あるいは, 全体の不均質性 I_t^2 を以下のように表現してみたくならない(Lim et al., 2014):

$$(3.13) \quad I_t^2 = \frac{\sigma_t^2 - \sigma_m^2}{\sigma_t^2}.$$

この指数は従来の I^2 と比較可能なものであり, その理由はどちらも分散の割合を示していて, サンプリングエラー σ_m^2 の影響を受けていないからである. これらの I^2 指数は, 一種の分散コンポーネントの解析の一種である. このような解析によって, どのレベルに大きな分散があるのかといったことが判明し, 次の節で説明するメタ回帰に利用できる.

4. メタ回帰モデル

一次研究間の分散がゼロより大きいことが確認された場合, ランダム効果モデルの不均質性を説明するために, メタ回帰(meta-regression)解析を実施しなければならない. メタ回帰は重みつき回帰の特殊な場合といえる. このことからわかるように, メタ回帰では予測変数(説明変数)が使用可能であり, 調整変数(moderator)と呼ばれている. 生態学・進化生物学のメタ解析で扱うデータは不均質であることが多く, メタ回帰はたいいていの場合において必要となる(Nakagawa and Santos, 2012; Mengersen et al., 2013). 以下では, まずこれまでのメタ回帰を紹介し, それにつづいて, 生態学・進化生物学のデータセットを解析するのに適した, メタ回帰のマルチレベルモデルを説明する.

4.1 基本となるメタ回帰モデル

メタ回帰モデルはランダム効果モデル(式(2.2))を改良したもので, 以下のように書ける:

$$(4.1) \quad \begin{aligned} z_j &= \eta_j + u_j + m_j, \\ \eta_j &= \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \beta_3 x_{3j} + \dots, \end{aligned}$$

すでに何度も登場した変数であるが、 u_j は一次研究 j に固有な効果、 m_j は一次研究 j におけるサンプリング誤差である ($j = 1, \dots, N_{\text{study}}$ であることに注意)。新しく追加されたものとしては、以下のとおりである： β_0 は切片(メタ解析の全平均)、 β_{1-3} は傾き(回帰係数)、 x_{1-3} は調整変数である。調整変数は2水準以上のカテゴリカル変数としてもよいが、その場合は3水準以上となるときにはダミー変数を使うことになるだろう。

メタ回帰の第一の目的は、観察された不均質性(ランダム効果モデル内の一次研究間のばらつき)をうまく説明できるような、調整変数の組みあわせを探し出すことである。この指標となるものとして、以下のように R^2 を定義してみよう (Aloe et al., 2010; Cheung, 2014)：

$$(4.2) \quad R^2 = 1 - \frac{\hat{\sigma}_{u1}^2}{\hat{\sigma}_{u0}^2},$$

ここで $\hat{\sigma}_{u0}^2$ と $\hat{\sigma}_{u1}^2$ はそれぞれ調整変数をいれた場合・いれなかった場合の一次研究間の分散の推定量である。このメタ解析の R^2 には(一次研究内の)サンプリングエラー σ_m^2 が含まれていないことに注意してほしい。なぜなら σ_m^2 は既知であり、説明されなければならない分散からは除外されているからである。この R^2 は便利な指標であるが、調整変数(予測子)の最良の組み合わせを選ぶための指標ではない。その理由は、調整変数を増やせば増やすほど R^2 が大きくなるからである。このため、たくさんの調整変数を使う場合には、赤池の情報量規準あるいは他の規準を使って、モデル選択をするのがよいだろう (Nakagawa and Santos, 2012)。モデル選択については、ここで説明するには大きすぎる話題なので、興味のある読者は Burnham and Anderson (2002) や Claeskens and Hjort (2009) を参照してほしい。

4.2 マルチレベルメタ回帰モデル

メタ回帰モデル(式(4.1))は、ランダム効果メタ解析モデルで説明したいろいろな問題をそのまま受けついでいるので、生態学・進化生物学での利用においても利用が制限されるかもしれない。マルチレベルメタ解析モデル(式(3.10))と同じように、このモデルも拡張してみよう。マルチレベルのメタ回帰モデルは、このように書ける：

$$(4.3) \quad \begin{aligned} z_i &= \eta_i + a_{k[i]} + w_{k[i]} + u_{j[i]} + e_i + m_i, \\ \eta_i &= \beta_0 + \beta_1 x_{1k[i]} + \beta_2 x_{2j[i]} + \beta_3 x_{3i}, \end{aligned}$$

ここで z_j は第 j 研究における効果量、 $a_{k[i]}$ は生物種 k の系統学的な効果、 $w_{k[i]}$ は系統学的な効果とは無関係な生物種 k 固有な効果、 $u_{j[i]}$ は一次研究 j に固有な効果、 e_i は第 i 効果量に固有な(それぞれの一次研究内の)ばらつき、 m_i は一次研究 i におけるサンプリング誤差、 x_1 は種レベルの調整変数、 x_2 は一次研究レベルの調整変数、 x_3 は効果量レベルでの調整変数 ($i = 1, \dots, N_{\text{effect}}$; $j = 1, \dots, N_{\text{study}}$; $k = 1, \dots, N_{\text{species}}$) である。調整変数の層に注目してもらうため、ここではわざと調整変数のレベルを変えてみた。上の式をみればわかるが、どのレベルの調整変数が有用であるかは、どの層がもっとも不均質であるかに依存している。たとえば、一次研究間レベルでの不均質性が一番大きい(つまり I_u^2 が大きい)のであれば、説明変数 x_2 を使うのが最も良いということになる。

すでに述べたように、メタ回帰解析のもっとも重要な目的は不均質性の説明であり、 R^2 はもっとも便利かつよく使われている指標であろう。マルチレベル版の R^2 は次のように定義できる：

表 1. この記事に登場するメタ解析モデル.

モデル	非独立性への対応	不均質性 (I^2) の使用	決定係数 (R^2) の使用	本文中の数式
固定効果メタ解析	No	No	No	式 (2.1)
ランダム効果メタ解析	No	Yes	No	式 (2.2)
階層メタ解析	Yes	Yes	No	式 (3.1), (3.2)
ランダム効果メタ回帰	No	No	Yes	式 (4.1)
階層メタ回帰	Yes	No	Yes	式 (4.3)

$$(4.4) \quad R^2 = 1 - \frac{\hat{\sigma}_{t1}^2 - \hat{\sigma}_m^2}{\hat{\sigma}_{t0}^2 - \hat{\sigma}_m^2},$$

ここで $\hat{\sigma}_{t0}^2$ と $\hat{\sigma}_{t1}^2$ は調整変数あり・なしの場合の全分散であり, $\hat{\sigma}_m^2$ は式(2.3)のサンプリングエラーに相当する (Nakagawa and Schielzeth, 2013). Cheung (2014) は各レベルにおける R^2 を定量すべきだと勧めていて, これはもともと Raudenbush (2009) がマルチレベルモデルで一般について述べていた. たとえば, 式(3.10)のモデルにおいて, 生物種固有なレベルの R^2 は $R_w^2 = 1 - \hat{\sigma}_{w1}^2 / \hat{\sigma}_{w0}^2$ と表現することができる. しかしながら, このようなアプローチは有効かもしれないが, R^2 を上のように定義してしまうと負の値になることもあり, 解釈ができなくなる (Nakagawa and Schielzeth, 2013). そのような理由から, マルチレベルメタ回帰モデルの R^2 (式(4.4)) を使ったほうが便利かもしれない. 以上で紹介したメタ解析のための統計モデルの特徴を, 表 1 にまとめて示した.

5. 興味あるトピックをいくつか

この総説では, メタ解析・メタ回帰についての統計学的な技術にまつわるすべての重要なトピックについて網羅しきれなかった. この節では, 4 個のトピックについて簡単に紹介してみたい. これらによっていくらかでも論点をおぎなえるだろう.

5.1 ソフトウェアの実装

生態学・進化生物学のメタ解析において, 従来型のランダム効果モデル(式(2.2))の利用が普及したのは, *MetaWin* (Rosenberg et al., 2000) のようなメタ解析の実装が使えるようになったことが原因だろう. ただし, このソフトウェアではマルチレベルメタ解析はできない(メタ解析の計算機プログラムの解説は Schmid et al., 2013 参照). しかしながら, 最近では, この総説で説明したすべての統計モデルを簡単に実装可能なフリーのソフトウェアが利用できるようになっている. とくに次の 2 つの R package は知っておくと良いだろう: i) *metafor* (Viechtbauer, 2010), ii) *MCMCglmm* (Hadfield, 2010). これら 2 つの大きなちがいは, 前者はパラメータの値を最尤法(頻度主義的な枠組)で推定し, 後者は名前からわかるように Markov chain Monte Carlo (MCMC) による推定をするベイズ的な枠組である. これらに加えて, *ASReml* (Gilmour et al., 2002) と *BUGS* (Lunn et al., 2000) はここで説明したすべてのモデルを実装できるし, 他にもより多くのソフトウェアが見つかるだろうし新しく作られていくだろう.

5.2 論文公表バイアスの解析

メタ解析における論文公表バイアスはかなり重要であり, この問題だけを扱う書籍が出版されているほどである (Rothstein et al., 2005). すべてのメタ解析モデルの基本的な仮定は, 論文

公表バイアスはないということである。つまり、メタ解析で使っている効果量はバイアスのないサンプルで構成されていることになっている。しかしながら、「有意ではない」結果となった研究より、「有意である」と結論している研究のほうが発表しやすいという論文公表バイアスが存在するのであれば、この「バイアスなし」の仮定はなりたたない。このような論文公表バイアスを検出し、バイアスの大きさを推定し、補正する統計学的手法が開発されてきた(Rothstein et al., 2005)。しかしながら、これらの統計学的手法には限界があり、その理由はだれも「真の」データセットから実際にどれぐらいの公表されなかった研究が失われているのか知らないためである。最近になって Simonsohn et al. (2014) は論文公表のバイアスを見つける新しい方法を提案し、それは有意性の p 値の分布 (p 曲線) を使用する。この手法では、失われた研究の個数をまったく知らなくてもよいとしている。シミュレーション研究によってこの手法の有効性が示されたとされる。この新しい方法が本当に有効かどうか注目していきたい。

5.3 欠測データ

論文公表バイアスの結果の研究消失は欠測データの一種と見ることができる。しかしながら、もし全ての研究データが入手できたとしても、欠測データの問題から逃れられるわけではない。さらに二種類の欠測データ、すなわち i) 効果量の欠測と ii) 調整変数の欠測にわけられる。これらの欠測が生じる原因は、主として論文著者が報告しないことによるものである。たとえば、研究方法と統計学的手法が明記されていないと、効果量の計算ができなくなったり、調整変数の情報が使えなくなる。Ellington et al. (2015) の最近の研究では、生態学・進化生物学のメタ回帰解析で調整変数の欠測データをあつかい、偏りが生じないようにパラメーターを推定するために、データ補完法を使う方法を説明している (Penone et al., 2014 も参照)。

5.4 多変量のメタ解析

生態学・進化生物学では多数の応答変数を同時に扱うような、多変量のメタ解析はめったにない。しかしながら一方で、医学・社会科学の分野では珍しいことではない。その中でもとくに、二組の効果量のセットを扱う二変量メタ解析モデルがもっともよく使われていて、多くのソフトウェア実装が利用できる (Mavridis and Salanti, 2013 で紹介されている)。このような対照的な状況ではあるが、生態学・進化生物学の分野でも多変量メタ解析が有用かもしれない。その理由は、2つあるいはそれ以上の個数の効果量の関係に興味がある場合もかなりあるからである。たとえば、すでに第1節で簡単に説明したように、Cleasby and Nakagawa (2012) は社会的一夫一妻制の鳥類の複数の種について、オスの齢とつがい内父性の関係、オスの齢とつがい外父性の関係、そして齢による父性の変化について調べた。最初の2つはいわゆる相関 (correlation) が効果量であり、「齢による父性の変化」の効果量は標準化した平均の差 (standardized mean difference, SMD) である。これらの統計量は Fisher の Z (transformation Z_r) に変換してメタ解析モデルをあてはめた。

6. おわりに

この総説では、生態学・進化生物学の分野におけるメタ解析・メタ回帰の統計モデルを概観してきた。現在においては、メタ解析・メタ回帰の柔軟な問題解決の枠組を使って、生態学・進化生物学分野で、よくあつかうタイプの相関構造 (非独立性) を解明できるだろう。しかしながら、このような目的にもっともかなったモデルであるマルチレベルのメタ解析・メタ回帰はまだそれほど普及していない。これはおそらくマルチレベル (かつ系統学的) メタ解析のソフトウェア実装がごく最近まで利用できなかったからだろう。そして、おどろくべきことなのか

しれないが, Koricheva and Gurevitch (2014)によると, 植物生態学の分野でのメタ解析では, 不均質性(heterogeneity)に関連する統計量さえ考慮されてもいないとのことである。現在の実態を考えると, 生態学・進化生物学の分野の研究者に対して, 早急にメタ解析の技法の教育が必要なかもしれない。実際のところ, 生態学・進化生物学の分野では, 学部生だけでなく大学院生に対する正規の統計学授業でもメタ解析が教えられていない。ここまで見てきたように, メタ解析とメタ回帰は線形モデルの単なる拡張にすぎない。しかしながら, たとえば I^2 のようなメタ解析に特有な観念を理解することは, メタ解析を実施するだけでなく, 文献中に登場するメタ解析の結果を解釈するためにも欠かせない。この強力な統計学的手法は重要であり, ますます広く使われることになるだろうから, 生態学・進化生物学のメタ解析の適切な訓練プログラム作りが喫緊の課題だといえる。

謝 辞

初期の原稿にコメントをくれた Losia Lagisz に感謝したい。また詳しく査読し改定案を提示してくれた査読者・編集者たちにも深く感謝したい。この研究について, 著者の一人である中川震一は ARC Future Fellowship(FT130100268)から援助していただいた。

参 考 文 献

- Adams, D. C. (2008). Phylogenetic meta-analysis, *Evolution*, **62**(3), 567–572.
- Aloe, A. M., Becker, B. J. and Pigott, T. D. (2010). An alternative to r^2 for assessing linear models of effect size, *Research Synthesis Methods*, **1**(3-4), 272–283.
- Arnqvist, G. and Wooster, D. (1995). Meta-analysis: Synthesizing research findings in ecology and evolution, *Trends in Ecology & Evolution*, **10**, 236–240.
- Bininda-Emonds, O. R. P., Cardillo, M., Jones, K. E., MacPhee, R. D. E., Beck, R. M. D., Grenyer, R., Price, S. A., Vos, R. A., Gittleman, J. L. and Purvis, A. (2007). The delayed rise of present-day mammals, *Nature*, **446**(7135), 507–512.
- Borenstein, M. (2009). Effect sizes for continuous data, *The Handbook of Research Synthesis and Meta-analysis* (eds. H. Cooper, L. V. Hedges and J. C. Valentine), 221–235, Russell Sage Foundation, New York.
- Burnham, K. P. and Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-theoretic Approach*, 2nd ed., Springer-Verlag, Berlin.
- Chamberlain, S. A., Hovick, S. M., Dibble, C. J., Rasmussen, N. L., Van Allen, B. G., Maitner, B. S., Ahern, J. R., Bell-Dereske, L. P., Roy, C. L., Meza-Lopez, M., Carrillo, J., Siemann, E., Lajeunesse, M. J. and Whitney, K. D. (2012). Does phylogeny matter? Assessing the impact of phylogenetic information in ecological meta-analysis, *Ecology Letters*, **15**(6), 627–636.
- Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach, *Psychological Methods*, **19**(2), 211–229.
- Claeskens, G. and Hjort, N. L. (2009). *Model Selection and Model Averaging*, Cambridge University Press, Cambridge.
- Cleasby, I. R. and Nakagawa, S. (2012). The influence of male age on within-pair and extra-pair paternity in passerines, *Ibis*, **154**(2), 318–324.
- Cochran, W. G. (1954). The combination of estimates from different experiments, *Biometrics*, **10**(1), 101–129.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Hillsdale, New Jersey.

- Cooper, H., Hedges, L. V. and Valentine, J. C. (2009). *The Handbook of Research Synthesis and Meta-analysis*, 2nd ed., Russell Sage Foundation, New York.
- Cornwallis, C. K., West, S. A., Davis, K. E. and Griffin, A. S. (2010). Promiscuity and the evolutionary transition to complex societies, *Nature*, **466**(7309), 969–972.
- Curtis, P. S. and Queenborough, S. A. (2012). Raising the standards for ecological meta-analyses, *New Phytologist*, **195**(2), 279–281.
- Egger, M., Smith, G. D. and Altman, D. G. (2001). *Systematic Reviews in Health Care: Meta-analysis in Context*, 2nd ed., BMJ, London.
- Ellington, E. H., Bastille-Rousseau, G., Austin, C., Landolt, K. N., Pond, B. A., Rees, E. E., Robar, N. and Murray, D. L. (2015). Using multiple imputation to estimate missing data in meta-regression, *Methods in Ecology and Evolution*, **6**(2), 153–163.
- Evans, S. R., Hinks, A. E., Wilkin, T. A. and Sheldon, B. C. (2010). Age, sex and beauty: Methodological dependence of age- and sex-dichromatism in the great tit *Parus major*, *Biological Journal of the Linnean Society*, **101**(4), 777–796.
- Garamszegi, L. Z. (ed.) (2014). *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology*, Springer, New York.
- Gelman, A. and Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, Cambridge.
- Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J. and Thompson, R. (2002). *ASReml User Guide Release 1.0.*, VSN International Ltd, Hemel Hempstead, UK.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research, *Educational Researcher*, **5**, 3–8.
- Glass, G. V. (2015). Meta-analysis at middle age: A personal history, *Research Synthesis Methods*, **6**, p.2.
- Gleser, L. J. and Olkin, I. (2009). Stochastically dependent effect sizes, *The Handbook of Research Synthesis and Meta-analysis* (eds. H. Cooper, L. V. Hedges and J. C. Valentine), Russell Sage Foundation, New York.
- Gurevitch, J. and Hedges, L. V. (1999). Statistical issues in ecological meta-analyses, *Ecology*, **80**(4), 1142–1149.
- Hadfield, J. D. (2010). MCMC methods for multi-response generalised linear mixed models: The MCMCglmm R package, *Journal of Statistical Software*, **33**, 1–22.
- Hadfield, J. and Nakagawa, S. (2010). General quantitative genetic methods for comparative biology: Phylogenies, taxonomies and multi-trait models for continuous and categorical characters, *Journal of Evolutionary Biology*, **23**, 494–508.
- Harvey, P. H. and Pagel, M. D. (1991). *The Comparative Method in Evolutionary Biology*, Oxford University Press, Oxford.
- Hedges, L. and Olkin, I. (1985). *Statistical Methods for Meta-analysis*, Academic Press, New York.
- Hedges, L. V. and Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis, *Psychological Methods*, **3**(4), 486–504.
- Higgins, J. P. T. and Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis, *Statistics in Medicine*, **21**(11), 1539–1558.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J. and Altman, D. G. (2003). Measuring inconsistency in meta-analyses, *British Medical Journal*, **327**(7414), 557–560.
- Higgins, J. P. T., Thompson, S. G. and Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis, *Journal of the Royal Statistical Society Series A-Statistics in Society*, **172**, 137–159.
- Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. and Mooers, A. O. (2012). The global diversity of birds in space and time, *Nature*, **491**(7424), 444–448.

- Kamiya, T., O'Dwyer, K., Nakagawa, S. and Poulin, R. (2014). What determines species richness of parasitic organisms? A meta-analysis across animal, plant and fungal hosts, *Biological Reviews*, **89**(1), 123–134.
- Koricheva, J. and Gurevitch, J. (2014). Uses and misuses of meta-analysis in plant ecology, *Journal of Ecology*, **102**(4), 828–844.
- Koricheva, J., Gurevitch, J. and Mengersen, K. (2013). *The Handbook of Meta-analysis in Ecology and Evolution*, Princeton University Press, Princeton.
- Lagisz, M., Hector, K. L. and Nakagawa, S. (2013). Life extension after heat shock exposure: Assessing meta-analytic evidence for hormesis, *Ageing Research Reviews*, **12**(2), 653–660.
- Lajeunesse, M. J. (2009). Meta-analysis and the comparative phylogenetic method, *American Naturalist*, **174**(3), 369–381.
- Lajeunesse, M. J. (2011). On the meta-analysis of response ratios for studies with correlated and multi-group designs, *Ecology*, **92**(11), 2049–2055.
- Liermann, M. and Hilborn, R. (1997). Depensation in fish stocks: A hierarchic Bayesian meta-analysis, *Canadian Journal of Fisheries and Aquatic Sciences*, **54**(9), 1976–1984.
- Lim, J. N., Senior, A. M. and Nakagawa, S. (2014). Heterogeneity in individual quality and reproductive trade-offs within species, *Evolution*, **68**(8), 2306–2318.
- Lunn, D. J., Thomas, A., Best, N. and Spiegelhalter, D. (2000). Winbugs—A Bayesian modelling framework: Concepts, structure and extensibility, *Statistics and Computing*, **10**(4), 325–337.
- Mavridis, D. and Salanti, G. (2013). A practical introduction to multivariate meta-analysis, *Statistical Methods in Medical Research*, **22**(2), 133–158.
- Mengersen, K., Jennions, M. D. and Schmid, C. H. (2013). Statistical models for the meta-analysis of nonindependent data, *The Handbook of Meta-analysis in Ecology and Evolution* (eds. J. Koricheva, J. Gurevitch and K. Mengersen), Chapter 16, 255–283, Princeton University Press, Princeton.
- Nakagawa, S. and Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists, *Biological Reviews*, **82**(4), 591–605.
- Nakagawa, S. and Poulin, R. (2012). Meta-analytic insights into evolutionary ecology: An introduction and synthesis, *Evolutionary Ecology*, **26**(5), 1085–1099.
- Nakagawa, S. and Santos, E. S. A. (2012). Methodological issues and advances in biological meta-analysis, *Evolutionary Ecology*, **26**(5), 1253–1274.
- Nakagawa, S. and Schielzeth, H. (2010). Repeatability for gaussian and non-gaussian data: A practical guide for biologists, *Biological Reviews*, **85**(4), 935–956.
- Nakagawa, S. and Schielzeth, H. (2013). A general and simple method for obtaining r^2 from generalized linear mixed-effects models, *Methods in Ecology and Evolution*, **4**(2), 133–142.
- Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., Young, B. E., Graham, C. H. and Costa, G. C. (2014). Imputation of missing data in life-history trait datasets: Which approach performs the best?, *Methods in Ecology and Evolution*, **5**(9), 961–970.
- Raudenbush, S. W. (2009). Adaptive centering with random effects: An alternative to the fixed effects model for studying time-varying treatments in school settings, *Education*, **4**(4), 468–491.
- Raudenbush, S. W. and Bryk, A. S. (2002). *Hierarchical Linear Models: Application and Data Analysis Methods*, Sage, Thousand Oaks.
- Rosenberg, M. S., Adams, D. C. and Gurevitch, J. (2000). *MetaWin: Statistical Software for Meta-analysis*, 2nd ed., Sinauer, Sunderland, Massachusetts.
- Rothstein, H., Sutton, A. J. and Borenstein, M. (eds.) (2005). *Publication Bias in Meta-analysis: Prevention, Assessment and Adjustments*, John Wiley, Chichester.
- Sánchez-Meca, J. and Marín-Martínez, F. (2008). Confidence intervals for the overall effect size in

- random-effects meta-analysis, *Psychological Methods*, **13**(1), 31–48.
- Schmid, C. H., Stewart, G. B., Rothstein, H. R., Lajeunesse, M. J. and Gurevitch, J. (2013). Software for statistical meta-analysis, *The Handbook of Meta-analysis in Ecology and Evolution* (eds. J. Koricheva, J. Gurevitch and K. Mengersen), Chapter 12, 174–191, Princeton University Press, Princeton.
- Simonsohn, U., Nelson, L. D. and Simmons, J. P. (2014). P-curve and effect size: Correcting for publication bias using only significant results, *Perspectives on Psychological Science*, **9**(6), 666–681.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package, *Journal of Statistical Software*, **36**, 1–48.
- Weir, L. K., Grant, J. W. A. and Hutchings, J. A. (2011). The influence of operational sex ratio on the intensity of competition for mates, *American Naturalist*, **177**(2), 167–176.

Statistical Models for Meta-analysis in Ecology and Evolution

Shinichi Nakagawa¹ and Takuya Kubo²¹Evolution & Ecology Research Centre and School of Biological, Earth and Environmental Sciences,
University of New South Wales²Graduate School of Environmental Earth Science, Hokkaido University

Meta-analysis is now the gold standard for quantitatively summarizing primary studies not only in medical and social sciences, but also in the field of ecology and evolution. Meta-analytic techniques have primarily been developed in medical and social sciences where data sets (i.e. a collection of effect sizes) for meta-analysis are likely more homogeneous, and the data probably have less inter-dependence than those in ecology and evolution. Perhaps, because of this history, two original models of meta-analysis are not actually suitable for modeling data sets from the field of ecology and evolution; the two models are known as the fixed-effect and random effects meta-analysis, both of which assume independence among effect sizes. Meta-analyses in ecology and evolution often need to deal with two types of dependence (or correlated structures) in data: 1) dependent effect sizes within studies, and 2) dependence due to phylogenetic relatedness among species. We review statistical models of meta-analysis, which have been proposed to resolve these two types of dependence. We show that multilevel modeling incorporating phylogenetic comparative methods (termed ‘phylogenetic multilevel meta-analysis’) can appropriately handle typical meta-analytic data in ecology and evolution. We also discuss the concepts of heterogeneity (as I^2) in meta-analysis and of R^2 in meta-regression analysis. Although statistical models suitable for ecological and evolutionary meta-analyses are now available, the use of such models are still limited. Effective educational programs are now required to introduce these suitable meta-analytic models to ecologists and evolutionary biologists.