Kay, William Penry, Sutherland, Chris and Marques, Tiago Andre 2024. Training statistics-savvy ecologists. A call to action for improved statistical education in ecology. The Wildlife Professional 18 (6) , pp. 55-59.

Training statistics-savvy ecologists

A Call to Action for Improved Statistical Education in Ecology

By William P. Kay, Chris Sutherland and Tiago A. Marques

For a budding ecologist, comprehending vast datasets—temperature readings, population counts and ecosystem measurements—may seem insurmountable without the right tools and training.

Ecologists understanding data is akin to detectives solving a mystery. Just as Sherlock Holmes needs a magnifying glass to conduct his work, ecologists need statistics to address important crises, from devastating biodiversity loss to catastrophic climate change.

Ecological systems are inherently variable, and ecologists typically rely on observational data, where few—if any—sources of variation can be controlled. Statistics helps them navigate that unpredictability and untangle the signal—or lack thereof—from the noise. Statistics uncovers hidden patterns in all-too-often messy data, revealing the subtle threads that connect seemingly random events.

As technological advances abound, methods to analyze increasingly vast and complex data follow (Gilbert et al. 2024). In the era of big data in biology—where large datasets are generated from citizen science projects, genomics, wildlife tracking and epidemiological research—statistics is essential to interpreting findings, linking results to underlying biological mechanisms and making predictions. However, many ecologists' ability and confidence to use novel data streams and apply new—or old—methods lag behind, leading to the incorrect use of statistics, including a reliance on overly simplistic, and potentially inappropriate, statistical tools.

As statistics educators, we feel that much of this challenge lies in the misunderstandings and limitations of teaching statistics to early-career scientists.

Statistical struggles

The wildlife field widely recognizes that statistics is crucial for ecological research. Popular statistical software packages, such as R, have been specifically adapted for wildlife biology, and there's been a proliferation of accessible 'introductory statistics for nonstatisticians' books. But challenges in training statistically literate ecologists have existed for some time and still remain.

One reason for this is statistics anxiety—the apprehension that a student experiences when exposed to statistics (Onwuegbuzie & Wilson, 2003)—which is compounded by the technical vocabulary and lexical abstractness of statistics (Rollings, 2024). This phenomenon extends beyond formal education. Many biologists view the statistical training opportunities afforded to them, as well as their own abilities, as inadequate (Barone et al. 2017; Attwood et al. 2019; Williams et al, 2019).

But a broader challenge lies in university curricula. Ecology programs lack standardization in their statistics training at both undergraduate and postgraduate levels. As a result, some courses may focus too narrowly on the use of specific tools or methods—like t-tests or correlation analyses—and miss a broader appreciation of foundational statistical literacy and intuition—like randomness and variability. This can hinder students' ability to identify solutions to novel problems, which for anything but the simplest cases, are in practice often not amenable to existing "off-the-shelf" solutions and require proper (statistical) thought.

However, challenges like time restrictions, heavy workloads and low staff capacity result in a gap between what quantitative educators want to teach and the capacity they have to do so. In addition, while statistics teachers may be experts at statistics, they may not be experts at *teaching* statistics (Zieffler et al. 2008). This, together with a lack of clear direction in curricula, can push educators toward "folk pedagogy"—relying on personal experience and preference rather than evidence-based practice. Statistics classes that would benefit from similar approaches are instead highly variable in structure, delivery and assessment methods, and contact hours. As a result, ecologists are all getting different training in statistics.

We have yet to see widespread implementation of effective solutions. But it's clear we need a systemic shift across biological disciplines to improve statistics training. And among ecology and the wider life sciences, there is appetite and a need for change. So, what should we do?

Curriculum development

There is no silver-bullet solution to these challenges, but we are optimistic about the potential for progress. We offer some suggestions for closing the statistics education gap in ecology and wildlife biology.

An essential first step is to develop an agreed-upon standardized statistics curriculum for ecology programs with a core set of learning outcomes and foundational concepts. However, we know that every class can't be exactly the same, and we recognize the value of allowing instructors flexibility in their approaches.

There are already examples of what a standardized statistics curriculum could and should look like for biologists, which suggest the prioritization of training relating to experimental design, data exploration, statistical inference and critical appraisal of results (*e.g.*, Ellison & Dennis, 2010; Mirra & Thomas, 2023). For this to be achieved, educators need sufficient time allotted to teach statistics to biologists, something that is lacking at many institutions (Barraquand et al. 2014).

Educators are often not afforded enough time because institutions may undervalue statistics as a key skill. It wouldn't necessarily mean that every ecology degree needs to expand the amount of statistics training offered—though this may be necessary for some programs. However, many institutions should consider providing more time for students to do the statistics training currently expected of them *i.e.*, build in more consolidation time to deepen their understanding, have more opportunity to practice, and learn at a slower, more manageable pace.

Initial classes should focus on fundamental statistical concepts—randomness, variability, randomization, confounding, and the importance of data exploration, including visualization. This would help students develop universally applicable, foundational knowledge. Only after students have that foundation should they learn specific statistical methods.

Rather than having students apply what they learned once, a spiral curriculum—one that facilitates revisiting and continuing to apply statistical concepts and methods—could be used. There also needs to be a shift away from assessing rote memorization of statistical tests, and toward assessing students' ability to apply their skills of exploring and visualizing data, applying appropriate analytical tools to solve problems, and interpreting statistical results. These skills are essential for students to maintain proficiency and build confidence.

The use of simulations—and simulated data—is crucial to cementing understanding. However, to motivate learning, demonstrate relevance and mitigate anxiety, teachers should focus on solving real-world ecological problems using real-world data. In our experience, students understand the importance of statistics and become better statistical thinkers when they formulate simple ecological questions, collect data to answer them, and think about how statistics can be used. Students can benefit from hands-on applications of what they've learned through project work and laboratory and field practicals.

Getting with the program

The way quantitative material is taught is also important. We feel strongly that applied statistics is taught through the medium of R, a free, open-source software that allows users to perform graphical and statistical analysis. This software is popular among ecologists, and employers often seek applicants who are familiar with using it. Standardization to this platform would enable statistics teachers to better support students and each other.

Because computer programming goes hand-in-hand with statistics, as it promotes reproducibility and customization in analysis, it is also something students should learn. Programming could be taught separately from statistics, or it could be integrated and taught in parallel. There is merit in both approaches. Both skills can be challenging, and teaching both simultaneously can steepen the learning curve. For some students, learning programming can detract from learning statistics. As a result, either having programming 'boot camps' to provide dedicated training in programming or ensuring statistics courses have time dedicated to programming will provide the necessary scaffolding to ensure these learners can focus on building statistical confidence.

A scaffolded and strategic approach to statistics teaching will indirectly mitigate statistics anxiety. But there are also several ways to tackle this barrier directly. Addressing statistics anxiety from the start of a program is crucial, and regular support can help reduce it. One way to offer support is through drop-in clinics, which can provide a safe space for students to seek help and build confidence, and through dedicated training workshops.

Sharing personal experiences of overcoming statistics anxiety can also show students they're not alone and can be successful (Watt, 2023). We often hear students say, "I'm bad at math." We need to support students in realizing that statistics is not the same as mathematics and that competence in statistics does not require innate ability but rather persistence, practice and faith in the learning process. More holistically, encouraging students to consciously adopt a growth mindset—to view the challenge of learning statistics as an opportunity to develop a key skill and exciting tool rather than a necessary evil—can foster a positive attitude (Cuddington et al. 2023). Teachers should encourage students to embrace failure and develop resilience.

To bring about these changes, universities need to provide faculty with the necessary resources—sufficient time to teach in the curriculum, dedicated staff training, and the opportunity to undertake educational research and scholarship—to teach statistics effectively in ecology programs. Institutions should also recognize the particular challenges that educators who are teaching statistics face—such as that quantitative courses typically receive lower student satisfaction scores—and support them accordingly. This also requires creating interdisciplinary networks between ecology and statistics (Carey et al. 2019). Examples of these networks already exist: The Centre for Biomathematics at Swansea University, the Centre for Research into Ecological and Environmental Modelling (CREEM) at the University of St Andrews, and the Biostatistics Network at Cardiff University can serve as examples. Finally, national funding agencies should prioritize awarding funds for quantitative training initiatives in ecology.

Driving solutions forward

Unfortunately, despite growing recognition of the need to improve statistical training, progress feels frustratingly slow. We perceive this to be due, at least in part, to a lack of standards, departmental resistance, insufficient time for training, and a perception among students that Biosciences is a nonquantitative discipline. Open dialogue is essential, but we urgently need a collective shift from conversation to action.

For too long, there has been substantial variation in formal statistical training for nonstatisticians. Standardization of statistical training across Biosciences curricula is essential. We need to develop a flexible consensus about the concepts, skills and accompanying educational approaches to statistics that should be included in ecology curricula. This would provide a shared resource for teachers and instructors around the world.

We believe organizations like the UK's National Centre for Statistical Ecology (NCSE) and The Wildlife Society's Biometrics Working Group can contribute to designing such curricula. Indeed, organizations, funding bodies and institutions must show stronger leadership in promoting, mandating and providing quantitative training that is seen globally as an essential, yet vulnerable skill and to recognize the challenges faced by practitioners. At the graduate level, large funding programs should do more to develop cross-cutting or centralized support for quantitative training aimed at improving statistical literacy for all.

We invite ongoing collaboration and the sharing of best practices to drive these solutions forward. By working together, we can empower future generations to tackle the complex challenges of our time. This collective effort will not only benefit ecology but also contribute to a broader scientific and even societal landscape where quantitative fluency is a cornerstone of discovery, innovation and ultimately effective citizenship in a data driven world. Let us begin building a future where statistics is not a source of anxiety but a springboard for scientific exploration and problem-solving in ecology.

References

Attwood, T.K., Blackford, S., Brazas, M.D., Davies, A. and Schneider, M.V., 2019. A global perspective on evolving bioinformatics and data science training needs. *Briefings in Bioinformatics*, *20*(2), pp.398-404.

Barone, L., Williams, J. and Micklos, D., 2017. Unmet needs for analyzing biological big data: A survey of 704 NSF principal investigators. *PLoS computational biology*, *13*(10), p.e1005755.

Barraquand, F., Ezard, T.H., Jørgensen, P.S., Zimmerman, N., Chamberlain, S., Salguero-Gomez, R., Curran, T.J. and Poisot, T., 2014. Lack of quantitative training among early-career ecologists: a survey of the problem and potential solutions. *PeerJ*, *2*, p.e285.

Carey, C.C., Ward, N.K., Farrell, K.J., Lofton, M.E., Krinos, A.I., McClure, R.P., Subratie, K.C., Figueiredo, R.J., Doubek, J.P., Hanson, P.C. and Papadopoulos, P., 2019. Enhancing collaboration between ecologists and computer scientists: lessons learned and recommendations forward. *Ecosphere*, *10*(5), p.e02753.

Cuddington, K., Abbott, K.C., Adler, F.R., Aydeniz, M., Dale, R., Gross, L.J., Hastings, A., Hobson, E.A., Karatayev, V.A., Killion, A. and Madamanchi, A., 2023. Challenges and opportunities to build quantitative self-confidence in biologists. *BioScience*, *73*(5), pp.364-375.

Ellison, A.M. and Dennis, B., 2010. Paths to statistical fluency for ecologists. *Frontiers in Ecology and the Environment*, *8*(7), pp.362-370.

Gilbert, N.A., Amaral, B.R., Smith, O.M., Williams, P.J., Ceyzyk, S., Ayebare, S., Davis, K.L., Leuenberger, W., Doser, J.W. and Zipkin, E.F., 2024. A century of statistical Ecology. *Ecology*, p.e4283.

Medeiros Mirra, R. and Thomas, R.J., 2023. Killing me softly with your stats teaching: how much stats is too much stats?. In *Teaching Biostatistics in Medicine and Allied Health Sciences* (pp. 197-209). Cham: Springer International Publishing.

Onwuegbuzie, A.J. and Wilson, V.A., 2003. Statistics Anxiety: Nature, etiology, antecedents, effects, and treatments--a comprehensive review of the literature. *Teaching in higher education*, *8*(2), pp.195-209.

Rollings, L. (2024). Talking Statistics: A reflection on some of the problems with statistical language. *MSOR Connections*, *22*(3), 5–9. https://doi.org/10.21100/msor.v22i3.1476

Watt, H.C., 2023. Teaching Conceptual Understanding of p-Values and of Confidence Intervals, Whilst Steering Away from Common Misinterpretation. In *Teaching Biostatistics in Medicine and Allied Health Sciences* (pp. 43-59). Cham: Springer International Publishing.

Williams, J.J., Drew, J.C., Galindo-Gonzalez, S., Robic, S., Dinsdale, E., Morgan, W.R., Triplett, E.W., Burnette III, J.M., Donovan, S.S., Fowlks, E.R. and Goodman, A.L., 2019. Barriers to integration of bioinformatics into undergraduate life sciences education: A national study of US life sciences faculty uncover significant barriers to integrating bioinformatics into undergraduate instruction. *PLoS One*, *14*(11), p.e0224288.

Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K. and Chang, B., 2008. What does research suggest about the teaching and learning of introductory statistics at the college level? A review of the literature. *Journal of Statistics Education*, *16*(2).