UKCOTS GLASGOW 2025

UKCOTS25

University of Glasgow, 24-25 June 2025

Book of Abstracts

This conference is kindly sponsored by
Timberlake Consultants

timberlake

UKCOTS 25 is supported by the Glasgow Convention Bureau

UKCOTS25 Abstracts

1 TABLE OF CONTENTS

2 Presentation Abstracts

2.1 Tuesday 24th June, 09:45 – 10:30, parallel sessions

2.1.1 Innovative active learning

A doubly robust approach to assess the effects of flipping a Statistics course

Author: Theresa Schmitz

Using puzzles to teach statistics and probability

Author: Michael Fletcher

For several years I have produced the puzzle page for the magazine Significance. The brief has been to entertain and educate the readers. All statistics educators are faced with the same challenge. i.e. How does one make statistics informative and at the same time engage one's audience? This interactive session will look at the challenges facing educators and, using problems, consider how the challenges can be met.

2.1.2 Teaching tricky topics

A virtual experiment to teach experimental design

Author: Charlotte Jones-Todd

Bayes news: teaching the Beta-Binomial using real and fake headlines

Author: Laurie Baker

Determining whether a headline comes from a credible news source or a satirical outlet is not just a test of media literacy, it’s a chance to explore Bayesian Statistics. In this activity, students construct a prior based on how well they think they can classify real headlines from CNN and satirical headlines from The Onion. Students explore the proportion they guess correctly using a beta-binomial distribution in a Bayesian framework. Through this process, students:

- Explore the concept of a prior by comparing their initial beliefs to the data,
- Investigate how different priors affect posterior distributions,
- Visualize priors and posteriors through plotting, and
- Calculate summary statistics (e.g., means and credible intervals) for both priors and posteriors.

This hands-on activity combines critical thinking with practical skills in Bayesian inference, providing students with a rich understanding of how the beta-binomial distribution is applied to real-world problems and helping students across disciplines build skills in Bayesian thinking.

2.1.3 PROFESSIONAL TRAINING

Trends and topics in Statistics thinking: a data-driven analysis

Author: Mark Andrews

Statistics training through short courses and continuing professional development (CPD) events has remained consistently popular over the past few decades. This demand is driven by the increasing availability of data, advancements in computational tools, and a growing need for data literacy across disciplines. In response, universities, research institutions, and private enterprises across the UK have actively provided such training. Using computational text analysis of archives containing training course descriptions spanning over two and half decades, we investigate the current state of statistics training provision in the UK and how it has evolved. Specifically, we examine the key topics in statistics and data science being taught, the primary software tools used, the target audiences and scientific disciplines served, and the assumed level of prior knowledge (e.g., beginner, advanced). Additionally, we explore shifts in these trends over time. This large-scale archival analysis offers a comprehensive view of the statistics and data science training landscape. Our findings can help training providers better understand their market, anticipate emerging needs, and tailor their offerings to meet evolving demands.

UCL’s experience of training medical statistician apprentices

Author: Andrew Embleton

The UCL Institute of Clinical Trials and Methodology (ICTM) is a centre of excellence for clinical trials, meta-analysis, and epidemiological studies. Part of UCL's Faculty of Population Health Sciences, ICTM includes four specialised units dedicated to advancing global health through integrated research and training via two post-graduate courses. The UK government-backed Medical Statistician apprenticeship standard focuses on providing statistical leadership to multi-disciplinary teams across various sectors, including pharmaceuticals, healthcare, and biotechnology. Apprentices will ensure medical research is designed, conducted, analysed, interpreted, and reported in a statistically valid manner. This involves interacting with regulatory bodies, ethics committees, senior management, and clinical teams. Key Apprentice development points are designing studies, monitoring progress, analysing data, and writing reports. The apprenticeship with UCL integrates an MSc in Statistics for Clinical Trials, with the

dissertation forming part of the end-point assessment alongside a presentation and questioning on the Apprentice's work-based development. The first cohort started in September 2025, and insights and experiences gained from the first year at ICTM will be shared to increase understanding and the awareness of the Apprenticeship route for teaching statistics.

Transforming classroom education through self-paced content

Author: Dean Langan

The Centre for Applied Statistics Courses (CASC) at University College London (UCL) has delivered statistics training to students and professionals for nearly 20 years. Responding to the growing demand for flexible, self-directed learning, CASC has adapted by converting courses into a self-paced format, featuring short videos and interactive exercises. This initiative has been evaluated using demographic data, completion rates, participant feedback, and financial metrics. The findings highlight that self-paced learning serves as a valuable complement to traditional teaching methods, with many learners even expressing a preference for this format. Feedback has been overwhelmingly positive, with praise for the clarity and thoroughness of course design. Furthermore, the external market presents a sustainable funding model, enabling continued development and offering flexibility for professionals balancing career and learning commitments. This talk will present the evaluation findings and share practical insights on designing, implementing, and sustaining self-paced courses to empower educators and institutions in transforming statistics education.

2.2 Tuesday 24th June, 11:00 – 12:15, Parallel Sessions

2.2.1 Research Meets Teaching

For a few dollars more

Author: Simon Harden

Published papers on the teaching of Statistics are diverse, from theoretical through to very practical. Some even use statistics to analyse their findings! This presentation is a review of recent papers from journals dedicated to the teaching of Statistics and also from those with a wider brief. Subjects covered include practical applicability of any findings, types of students who are being taught and quality of statistical analyses. Suggestions of gaps and areas for improvement are also made.

UKCOTS25 Abstracts

Ethical considerations in statistical education research

Author: Bruce Dunham

Research in statistical education often involves behavioural studies with students as volunteer subjects. However, academic statisticians are typically not experienced in conducting studies with human subjects and therefore may not be aware of certain ethical considerations. Here the major ethical issues with respect to statistical education research are described, such as informed consent, potential risks, privacy, and sharing of data. The session aims to support academics in gaining ethical approval at their institutions for research in statistical education where responses on students are to be collected.

Crossing borders: engaging with research papers in Statistics

Author: Ioanna Papatsouma

Engaging undergraduate students with research at an early stage can enhance critical thinking, analytical skills, and communication in statistics. In this collaborative project between Imperial College London and Nanyang Technological University (NTU) Singapore, first- and second-year Mathematics students participated in guided discussions of research papers. Students formed reading groups, critically analysed assigned research papers, and engaged in facilitated discussions to deepen their understanding of key statistical concepts. A highlight of the project was the opportunity for students to interact directly with the paper authors, who were staff members of both institutions, allowing for deeper insights into statistical research and academic discourse. Through this process, students developed interdisciplinary perspectives, improved their ability to interpret research findings, and gained confidence in engaging with academic literature. In this talk, we will share key insights from the project, student reflections on their experience, and recommendations for integrating research-based teaching into the undergraduate statistics curriculum.

2.2.2 CURRICULUM DESIGN in LIFE SCIENCES

Redesigning a Nutrition & Dietetics Statistics curriculum: a stakeholder survey

Author: Erin Stella Sullivan

A strategic educational initiative was proposed to co-design a new statistics curriculum for nutrition & dietetics students, incorporating input from key stakeholders. A survey was circulated to students and staff within the King's College London Department of Nutrition & Dietetics, as well as external academics and potential graduate employers. The survey included both closed and open-text questions exploring current practice, willingness to learn new skills, barriers, facilitators, and perceptions of essential statistical competencies. Descriptive

analysis was conducted in Stata. Responses were obtained from n=18 external & n=15 internal academics, n=3 graduate employers and n=26 internal students. SPSS & R were most commonly used by staff, whilst students primarily used SPSS and Excel, with a strong interest in learning R. Key teaching challenges included limited coding experience, time constraints and difficulties in conveying advanced statistical concepts. Students' main challenges were a lack of coding experience, guidance, and mentorship. Students preferred workshops, whereas academics favoured online tutorials. Employers prioritised data ethics, advanced statistical understanding, storytelling and visualisation over software proficiency. Bridging the gap between the perspectives of learners, teachers, and employers is crucial to developing a curriculum that meets all stakeholders' expectations.

How should we train the next generation of statistical ecologists?

Author: William Kay

Ensuring that future ecologists have strong statistical knowledge is essential for reliable ecological inferences, reproducibility and scientific rigour. However, what constitutes ideal statistical training for ecologists remains up for debate, with limited guidelines internationally. In this talk, we present findings from a survey of experts at the International Statistical Ecology Conference (ISEC) 2020 and two subsequent expert roundtable discussions held at ISEC 2024 and BES 2024. Specifically, using a mixed-methods approach comprising quantitative and thematic analysis, we share perspectives on what statistical concepts, tools, and software are considered essential to be taught for ecologists, as well highlighting key themes such as curriculum structure, teaching methods, barriers to effective training, and community-driven solutions. We also emphasise the common challenges identified, including limited teaching time, the balance of theory with application, and the role of software and programming in learning statistics. From synthesising insights from both the survey and discussions, we propose how to best equip the next generation of ecological researchers with the statistical skills they need.

Current perspectives in teaching statistics in the life sciences in the UK

Author: Crispin Jordan

What do educators currently prioritize in teaching statistics and study design in the Life Sciences across the UK? We present results from discussions at two workshops (Manchester 2024, Edinburgh 2025) for educators in Biomedical Science, Biology, Medicine and Psychology, that addressed current practice in teaching undergraduate statistics. The participants considered seven themes: (1) Why do you think teaching is done as it is done within your discipline (2) What challenges does effective teaching face and what have disciplines done to address those challenges? (3) As educators, what are the current learning objectives when teaching study design and analysis? (4) Why do disciplines differ in topics being taught? (5) Should learning objectives be updated, and should teaching content change? (6) Which topics do you think students struggle with? (7) What pros/ cons do we anticipate for using simulations to teach mathematical concepts? We present a synthesis of primary conclusions from our workshops. We highlight diversity among disciplines with respect to teaching goals and

mechanisms that appear to effectively drive differences in time (credits) allocated to teaching study design and analysis. Simultaneously, we identify similar challenges in teaching this subject among the four disciplines.

2.2.3 COLLABORATIVE TEACHING AND LEARNING

Planning and teaching a community-engaged Data Science course

Author: Laurie Baker

I will discuss the course design and my experience teaching a 3.5 week intensive Community-Engaged Data Science and 10-week Community-Engaged Data Science at two liberal arts universities in Maine. I will share details of the course, including learning objectives, the course syllabus and resources used, and advice for other teachers interested in creating similar courses at their home universities based on the challenges I encountered and lessons learned from teaching this course. In the course, Community-Engaged Data Science, students work in multidisciplinary teams on a research problem identified by a community partner. The course emphasizes putting knowledge into practice, including going beyond individual fields of study to solve real world problems and understand community partner needs. Together with community partners, students work to gain insight from data, building skills in reproducible analysis and collaboration, using R programming tools and techniques. The course is designed as an advanced course building on skills gained in previous R programming and data analysis courses. Throughout the course students develop their mathematical and programming skills as well as skills and traits valued by employers of STEM professionals, such as teamwork, reproducible analysis, effective communication, independent thinking, and problem solving. Students also build skills in project management, using agile methodologies and weekly meetings with community partners designed to foster co-development and iterative and incremental project delivery.

Virtual exchange in teaching Statistics and Data Science

Author: Serveh Sharifi

Virtual Exchange is a collaboration between instructors from different institutions, often in different countries, to design shared student activities. This talk introduces the integration of virtual exchange in Statistics and Data Science education, emphasising the benefits of intercultural student interaction and active learning in improving students' data analysis and communication skills. Two virtual exchange projects between the University of Florida and the University of Edinburgh are presented as examples, and instructor and student perspectives on the effectiveness of the practices are described.

Co-creating data science assessment rubrics

Author: Jenny Sexton and Pete Edwards

In this presentation, we will present an innovative piece of ongoing work to involve students in the wording and presentation of assessment rubrics within Data Science. While analytic rubrics are commonly used in other disciplines, Data science students may not encounter this type of assessment until they undertake a project or dissertation. Providing students with access to marking rubrics is a first step towards transparent assessments but in practice students find rubrics opaque and difficult to engage with. To ensure that the skills and competencies required are clearly explained within our assessment rubric, we are working in partnership with students while the module is being developed. Involving students in the development of the module aims to ensure that markers expectations and students' interpretation of the wording of the assessments is aligned and provide all students with a sense of ownership.

2.3 TUESDAY 24TH JUNE, 13:15 – 14:15, WORKSHOPS

Using Generative AI to create R shiny applications

Authors: Tom Goodale

Many statistical concepts are best explained through interactive applications. R Shiny is a powerful framework for creating such applications, and there are many applications available on the internet. This hands-on workshop shows how to use Generative AI to streamline the process of creating Shiny applications to demonstrate statistical concepts. Participants will learn Shiny fundamentals, use AI chatbots to simplify application development, and explore ways of deploying them.

Using large lectures to support learning of quantitative methods

Author: Julie Scott Jones

Large lectures can be a vital tool in delivering quantitative methods teaching en masse. However, large lectures are often seen as passive and anonymous educational environments. This workshop will demonstrate that lectures can be a key platform of any quantitative methods module, through the embracing of the model of the lecture as a "theatre" of the "spectacle and surprise". Students attend lectures with preconceived (often negative) attitudes towards quantitative methods; by using the element of "surprise" and "spectacle", students can be reoriented towards more positive attitudes towards the subject, making them potentially more likely to engage. Relatedly, community, engagement and playfulness can be built that supports students learning. Interactivity is central to this approach; students are participants and spectators. This approach goes beyond the "flipped classroom" to embrace the often neglected "theatre" element of the large lecture. This workshop will outline the model on which

this approach is based, including empirical evidence of its efficacy; it will present examples of how to create “spectacle” and “surprise” in large quantitative methods lectures; and workshop attendees will be invited to participate as students.

Active learning with code-alongs, peer-reviews and Kaggle in a basic ML course

Author: Linda Hartman and Dmytro Perepolkin

Teaching statistical learning is no small feat: students often arrive with little to no experience in data analysis or programming. How can we fast-track their data literacy while keeping them engaged? In this session, we’ll take you beyond theory and into the classroom with an interactive demo that gameplays our own teaching methods, immersing you in the classroom experience. We’ll showcase our approach to accelerating statistical learning through live code-along sessions and interactive quizzes. We’ll also introduce the "blue kryptonite" skill set of key competencies that empower students to grok core machine learning concepts effectively. Our broader curriculum integrates assignment peer-reviews and an in-class Kaggle competition to sustain engagement and deepen learning. Drawing from our experience teaching basic statistical learning at Lund University, we’ll share key insights, lessons learned, and future directions for making statistical education more interactive and impactful.

2.4 TUESDAY 24TH JUNE, 15:45 – 17:00, PARALLEL SESSIONS

2.4.1 AI VS LEARNING

Students’ perceptions and use of Gen Ai Statistics education: UK & Australia

Author: Tjun Hoh

A team of colleagues who support the learning of statistics in a number of different universities across the UK and Australia, have undertaken a survey to explore students’ current levels of awareness, usage, and experiences of Generative AI in supporting their learning, and specifically its potential as an aid for learning statistics. Data on over 600 students was collected by an online questionnaire during 2024 and 2025 and some of the initial overall findings are reported. The questions address aspects such as the awareness and use of Generative AI tools among students, students’ awareness of the capabilities of Generative AI with supporting their learning generally and with learning Statistics, the extent to which students are using Generative AI in this way or feel it could be useful in this way, students’ perceptions of trust and use of AI in learning and assessments, as well as understanding the guidance they have received from their institutions thus far and the guidance they feel they need. If time allows, we may also have an opportunity to explore how some of these perceptions and views relate to students’ perceptions of their own ability/anxiety/confidence with statistics and/or help seeking behaviours.

Statistics assessment, generative AI and evaluative judgement

Author: Amanda Shaker

The emergence of Generative Artificial Intelligence (GenAI) has caused many educators to re-think assessment in Higher Education. While GenAI can be a valuable learning tool, many would agree that there is also a risk that it can undermine student learning in certain contexts. As such, responsible use of GenAI, along with appropriate assessment design, is crucial. In this talk, we present a framework for assessment design that has been developed in order to address this issue. The assessment design incorporates evaluative judgement of GenAI outputs, and also leverages R/Exams for efficiency of question generation and marking. Pedagogical considerations will be discussed, along with presentation of practical examples showing how the framework can be implemented. While the examples are specific to Statistics and Data Science, the general framework and pedagogical considerations are relevant to most disciplines.

Reflections on the use of Gen AI in teaching statistics and data science

Author: Ozan Evkaya

As a result of recent advancements in generative AI, teaching and learning in HE institutions are prone to certain changes. Given the diverse range of Gen-AI tools like ChatGPT or its competitors, there is an ongoing debate on implementing such tools for teaching and learning activities. This study aims to highlight the Statistics and Data Analysis (DA) capabilities of ChatGPT, assessing its performance while considering its bottlenecks and ongoing development. In addition to repeated prompting experiments on predefined tasks, the use of knowledge base systems relying on advanced prompting is critically examined for further integration of these tools into modern statistics and data science education. Both personal reflections and the review of the recent works will be leveraged to enrich the discussion.

2.4.2 EQUALITY, DIVERSITY AND INCLUSION

Applying the QuantCrit method: implications of critical race theory to Statistics teaching

Author: Ric Crossman

In their book How To QuantCrit: Applying Critical Race Theory to Quantitative Data in Education (2024), Strunk and Castillo lay out how to consider statistical practice through the lens of Critical Race Theory. The case is made that statistics is not a "color-evasive" endeavour, and

that an awareness of systemic racism is essential at every stage of statistical practice. In this session, the basic arguments of the book will be presented, and possible routes for how these arguments can influence or even be included in our teaching will be discussed.

There are no simple answers to complex questions

Author: Sean McKusker

The UN Sustainable Development Goals feature prominently in the profiles of many HE Institutions, whether directly within the curricula or as part of the impact measures by which institutions are ranked, including a measure of the research outputs which address SDGs. If we are to "ensure that all learners acquire the knowledge and skills needed to promote sustainable development" (SDG 4.7), statistical literacy across the range of academic disciplines within HE, has an important part to play. In this context, it is important that statistical literacy is reconceptualised in ways which recognise the affordances of technology and the wealth of data available. It needs to allow more inclusive access to the interpretation of that data, rather than that mediated by statistical "experts" who disseminate their interpretations, often as unimpeachable "facts". The work presented here, demonstrates a novel, graphical interface which allows users, including students and researchers within HE to interact with authentic complex, non-linear, multivariate data, such as those which underpin the issues of global educational inequality, without the technical barrier of conventional statistical techniques. It calls for the publication of data at higher levels of disaggregation to address the "tension between accuracy and accessibility" (Higgins, 2020) in such analyses.

Barriers towards statistics education for students with learning difficulties

Author: Megan Barnard

The last several years have seen an increase in the number of students declaring a specific learning difficulty (SpLD) in Higher Education (Brunswick et al., 2024). We also know that whilst SpLD students in STEM subjects perceive themselves to have strengths, they do not feel adequately supported by the Higher Education system (Griffin & Pollak, 2009; Syharat et al., 2023). These effects may be especially prevalent in subjects such as statistics, which results in high anxiety for the neurotypical student population. Having a better understanding of how SpLDs may impact statistics anxiety, as well as attainment on statistics modules, could be key to understanding how we can better support students in this domain. This study will conduct a secondary analysis on the SMARVUS dataset (Terry et al., 2023), a large-scale dataset with data from over 12,000 University students. This presentation will discuss, based on this analysis, whether SpLDs predict different types of statistics anxiety and attainment on statistics modules, and whether the year of study exacerbates any of these effects. Recommendations for statistics education provision will be discussed based on these results, such as when as well as what support should be provided to students with an SpLD.

UKCOTS25 Abstracts

2.4.3 STUDENT ENGAGEMENT

Can we really measure engagement? What Moodle logs can (and can't) tell us

Author: Laura Johnston

Student engagement is a widely debated concept with no universal definition. It is typically divided into behavioural, emotional, and cognitive components. While Virtual Learning Environments (VLEs) generate vast amounts of data, most agree that it primarily reflects behavioural engagement, leaving open questions about what can truly be measured. Despite the ongoing debate, educators call for a straightforward, practical way to track engagement in online learning spaces. This talk examines an adapted engagement metric that quantifies student study patterns using Moodle logs. By tracking the frequency, immediacy, and diversity of interactions across a module's chapters, this approach provides a structured method to assess how students engage with course materials. However, validating such a metric is challenging, as few established benchmarks exist for measuring engagement in VLEs. The session will explore what this metric can and cannot tell us about student engagement, discuss the limitations of using VLE data alone, and suggest how VLE design can be optimised to generate meaningful data for future learning analytics research.

Technology enhanced interactive blended learning to improve student engagement

Author: Nadarajah Ramesh

The way students learn in Higher Education institutions have always been changing influenced by many factors to include students' learning habits, forms of delivery, technological advancements etc, to name a few. Nevertheless, one key underlying principle central to students learning is student engagement with learning activities. It is a well-known fact that student understanding and performance would be better when they engage well with lessons and take part in learning activities. These are essential to their skill development as well as to the successful delivery of the module. Student engagement with lessons has been a concern to academics in recent years, especially during the post-pandemic years. With the provision of module handbooks or lecture notes and tutorials in advance, along with the facility to access lecture recordings and tutorial solution after the lesson, students tend to underestimate the importance of attending timetabled lessons and engaging with the learning activities in the classroom. This talk describes how student engagement in lessons has been enhanced, using technology enhanced blended learning, in some of the final year statistics modules offered to maths students at Greenwich.

Exploring student profiles in statistical performance

Author: Paddy O'Connor

Understanding factors influencing students' success in statistics is crucial, as many psychology students struggle with the subject, affecting their academic confidence and career readiness. This study examines profiles of undergraduate psychology students learning statistics, focusing on attitudes, IT self-efficacy, and statistical performance. A sample of 175 first-year students completed assessments on attitudes towards statistics, academic delay of gratification, test anxiety, and computer self-efficacy, plus an end-of-year exam. Latent profile analysis identified two groups: "Motivated Performers" with higher positive attitudes towards statistics, greater computer self-efficacy, and lower statistical anxiety, and "Anxious Learners," who scored lower in these areas. Although Motivated Performers outperformed Anxious Learners on the exam, the difference was nonsignificant. Academic delay of gratification was the only significant predictor of performance, regardless of profile membership. These findings emphasize promoting positive attitudes and IT confidence in statistics and highlight the importance of self-regulatory academic behaviours in preparing students for statistics exams.

2.5 WEDNESDAY 25TH JUNE, 9:00 – 10:00, WORKSHOPS

Teaching responsible model development with Generative AI

Author: Jennifer Gaskell

As generative AI tools, such as ChatGPT, become integral to model development, a critical gap remains: how can we ensure these models are created responsibly, efficiently, and ethically? Many organizations lack clear guidelines and training, whilst universities struggle to teach best practices. The CREATE Project, being developed at the University of Glasgow, aims to address this by building a comprehensive platform offering guidance for practitioners and educators on responsible generative AI use. This workshop will give a brief overview of the key ideas of the CREATE Project and introduce participants to the framework, focusing on four main principles:

- Education: Recognising both opportunities and risks of generative AI.
- Development: Adopting efficient, effective practices for model creation.
- Safety: Understanding ethical and appropriate AI use.
- Reporting: Ensuring the process of developing models is transparent, trustworthy, and reliable.

In this workshop, we will provide an overview of the course and give an opportunity for attendees to work through the CREATE resources. Following some introductory talks, there will be a series of discussions and an opportunity for interactive feedback. Talks: Jake Lever (Computer Science, University of Glasgow), Vinny Davies, Jennifer Gaskell and Craig Alexander (Maths & Statistics, University of Glasgow).

UKCOTS25 Abstracts

An R/Quarto tutorial for writing open educational materials

Author: James Bartlett

In statistics and programming courses, you might rely on expensive print textbooks which quickly become outdated or not cover your specific curriculum. Open educational materials allow you to write free materials which directly target the needs of your students. For scholarship, they also have the added benefit for educators in being a valuable output to demonstrate your learning and teaching practice. In this workshop, I will provide a tutorial on writing online books using R/Quarto and hosting them using GitHub pages. The combination of R and Quarto is perfect for statistics and programming content as you can combine code, statistics output, and LaTeX equations. If you can bring a laptop and follow the preparation instructions in Chapter 1 (https://bartlettje.github.io/dissemination_quarto/), the aim of the workshop is for everyone to create and host your initial book.

Using pair programming in statistics education

Author: Kasia Banas

Pair programming is a collaboration technique widely used in the software industry: it involves two people working together on one programming task. One person is the driver, suggesting solutions and typing the code; the other person is the navigator, helping with problem-solving, spotting mistakes, and acting as the sounding board for the driver. After a short time, they swap roles. The use of pair programming in education is also increasing, and it has been shown to result in better quality code and more discussion between students than solo programming (Hawlitschek et al., 2022). The literature shows that it is particularly beneficial for students with lower prior programming skills (Goel & Kathuria, 2010). At the University of Edinburgh, we use pair programming extensively in statistics, programming and data science courses. In this interactive workshop, we will give you a taste of pair programming, so you can try it in your own programming and teaching practice. We will have tasks prepared in R, Python and Excel. You will be paired with another person and given a set of small challenges to solve together. You will practice both pair programming roles and you will get a chance to reflect on your experience.

2.6 Wednesday 25th June, 10:45 – 12:00, Parallel Sessions

2.6.1 Assessment

Assessing critical evaluation of data science studies in the age of LLMs

Author: David Sterratt

UKCOTS25 Abstracts

Until 2024, we assessed a learning outcome of critically evaluating data-driven methods and claims from case studies via a coursework in which students answered a set of questions at various levels on Bloom's taxonomy about a data science study, typically a short scientific article, and an accompanying media article. We believed, and students reported, that examining how statistical concepts are used in a real study and how they are reported in the media was a valuable experience. Large language models can now do the task well enough to get a more than passing grade. It is not currently possible to detect LLM-generated text with enough certainty to count as academic misconduct, and although the LLM-generated work has flaws, it is time-consuming for markers to identify them. We are therefore attempting to assess critical evaluation in an exam, while still retaining the experience of engaging with an authentic study. Our general approach is to introduce a study to students throughout the semester, and write exam questions related to the study. This talk will describe our approach and evaluate how well it has worked this year.

Assessment of online statistics modules

Author: Emma Howard

This talk is about my experience teaching two four-week long online statistics modules; Time Series and Multivariate Analysis. The modules are intended for graduates of disciplines, other than statistics, who want to develop their knowledge of statistical methods for solving problems. Both modules have a similar fusion format; the module content is taught through videos, and students have 1-2 hours per week of live online sessions. In 2023/24, the assessment for the modules consisted of formative quizzes, a project (40%), and an online end-of-semester examination (60%). In 2024/25, the assessment consists of a project (80%) and quizzes (20%). Focusing on Multivariate Analysis, students are asked to choose one of three datasets, set their own research questions, choose the variables needed to address their questions, analyse the dataset showing mastery of the module content, and write a cohesive report. Students can submit three project pages on their exploratory analysis in advance for formative feedback; which previous students have commented positively on. However, not all students choose to engage with the formative report, and plagiarism by students using GenAI tools is a concern. Through this talk, I seek to promote a discussion on the assessment of online statistics modules.

Interactive dashboards for design and analysis of experiments

Author: Rachel Oughton

Being an applied statistician often involves working through the design and analysis of an experiment (or trial, study etc.). Choices made at one stage will impact subsequent stages, and the statistician is required to understand the process as a whole. Assessing this can be difficult, particularly in a typical exam paper. I will discuss my use of Shiny dashboards for assessments in a 4th year Clinical Trials module. Students are given a “trial scenario” introducing the trial and some key features of the eligible population and outcome variables. The dashboard uses

an underlying “demographic” model to generate participant data given a sample size value. Once those participants have each been allocated to a trial arm, a “trial” model is used to generate outcome data for the student to analyse. Although the assessment is structured, it is also somewhat divergent in that students can at each stage choose which method they use. Students are to submit a written report, explaining their reasoning behind each step, communicating and interpreting results and commenting on the process.

2.6.2 ENGAGING NON-SPECIALIST STUDENTS

Engaging veterinary students with Statistics: evolving with curriculum changes

Author: Alice Batchelor and Liz Grant

How can we make statistics engaging for Veterinary Medicine students within a packed curriculum? At the University of Surrey, the Maths and Statistics Advice (MASA) service has worked with the School of Veterinary Medicine for several years to support with the statistics content of the two third-year Veterinary Research modules, through full-cohort teaching and 1:1 MASA drop-ins and appointments. The focus has shifted in the last year, following the new module lead's replacement of the 'Veterinary Research Project' with a 'Critical Appraisal, Data Interpretation, and Study Design' module, better aligning with skills relevant to veterinary practice. In this talk, we will share our approaches and how we adapted, shifting from software-based workshops on conducting and presenting statistics to content focused on choosing, interpreting, and critiquing them. Inspired by UKCOTS24, we now use live data, allowing students to see and engage with real-time applications of statistics, and we integrate research paper examples throughout to provide contextual learning. We also draw on our extensive 1:1 experience with students to inform our teaching. We'll reflect on the positive impact of the changes, while acknowledging the challenges that remain in facilitating engaging, effective statistical learning within the constraints of time and module structure.

Biology students and varied outcomes: the struggle to see the need for data analysis

Author: Daniel Franklin

Many students entering higher education (HE) biology courses struggle with statistical analysis: some due to weak mathematical skills, but others because they do not see its relevance. Pre-university Biology curricula often present biological processes as deterministic, emphasizing singular “correct” outcomes. Statistical concepts, where included, appear in isolation, rather than as integral to understanding biological variability. As a result, students arrive at university with a low tolerance for statistical analysis and an expectation that biology is a descriptive science, questioning why data analysis is necessary. Our previous research (Franklin & Harrison, 2022) found that undergraduate students did not feel adequately prepared for the

quantitative demands of a biology degree. Here, I will demonstrate how pre-university qualifications create this issue by failing to present biology as a quantitative discipline, overlooking that natural processes lead to varied outcomes: the fundamental premise of biological data analysis.

Many students only recognize the need for statistics late in their degree, often too late to fully embrace it. I will discuss strategies for fostering early buy-in and propose ways for pre-university curricula to better reflect biology's quantitative nature, reducing misconceptions and enabling them to use statistics to engage more deeply with the subject of biology.

2.6.3 NEW TRICKS FOR CLASSIC STATS

Students can design better summary tables – and you can too!

Author: Silas Bergen

Data summary graphs and tables are a crucial component of any report, poster, or presentation that discusses data analysis results. While teaching an undergraduate statistical consulting class a few years ago, I grew increasingly appalled at the quality of the graphs tables my students were including in their consulting reports. But what was it about their work that was so troublesome? In this talk, I will discuss how my students' subpar work forced me to think about the principles that govern high-quality graphs and tables, focusing primarily on optimal grouping arrangements (Gestalt principles). I will present these principles to participants and demonstrate how optimal use of these principles can improve the clarity of a summary table or graph. I hope participants will leave this presentation with a better understanding of how to create more communicative data summary tables and graphs and how to help their students do the same.

Three exercises for the "world beyond $p < 0.05$"

Author: Peter Martin

Misuse of statistical inference procedures is widespread. In place of thoughtful inference, a mindless "null ritual" holds sway that focuses on testing only null hypotheses of "no effect" and takes the 5 % threshold for p-values as the main or only standard by which to judge the strength of statistical evidence. The null ritual goes hand in hand with numerous misconceptions that encourage poor research design, poor analytic decisions, and misinterpretation of statistical results. Debates on how to move to "a world beyond $p < 0.05$" often emphasize the importance of statistics education, yet there are few published proposals on how to teach statistical inference in a way that combats pre-existing misconceptions, enables students to reason under conditions of uncertainty, and encourages thoughtful, open and appropriately modest conduct and interpretation of statistical inference. This presentation will show three exercises that aim to encourage critical thinking in data analysis. The exercises

directly address three common misconceptions about statistical inference in the context of regression modelling. The misconceptions relate to the selection of variables in regression models, the use of "statistical significance" to judge consistency of results across studies, and the nature of "null hypotheses" that should and can be tested.

2.7 Wednesday 25th June, 13:00 – 14:15, Parallel Sessions

2.7.1 Enthusing about basic statistics

A Trojan horse for teaching statistical thinking to Health Data Science students

Author: Jamie Sergeant

Health Data Science students may be more interested in learning about advanced modelling techniques than sound statistical thinking. In this postgraduate taught module, machine learning (ML) was used as a Trojan horse, attracting and gaining acceptance from students who then developed their statistical thinking skills through critical appraisal of published health research studies which used ML. Teaching was delivered as part of an optional 15-credit module on advanced statistical topics. The module aimed to foster statistical thinking, enable engagement with published research, and promote informed discussion and debate. An active learning approach was used, whereby students engaged with research through directed reading and enquiry as independent study. In-person classes were used for the reinforcement of core content, practical exercises and group presentations. Experiential and research-led learning saw students undertake their own risk of bias assessments on published research studies, first as a formative exercise, then as a summative assessment delivered as either an exam or an individual coursework assignment. Almost all students were able to competently perform a risk of bias assessment and adequately justify their decisions. Health Data Science students developed their statistical thinking skills through the authentic undertaking of critical appraisal and risk of bias assessment.

A Statistics course with a triple challenge

Author: Mark Van Lokeren

In this talk, we reflect on the design and delivery of an introductory course in statistics which is offered to postgraduate students. What the course makes stand out is that it faces a triple challenge:

(1) Students taking the course come from all sorts of background, ranging from complete beginners to people with considerable prior knowledge. Moreover, some students enter the

course straight from undergraduate studies while others return to studies after having worked for a couple of years in industry.

(2) Apart from the statistical content, the course also serves as an introduction to the programming language R. Coding experience varies substantially among students.

(3) The course is offered fully online as students join from different continents. Most learning takes place asynchronously, alternated with a few synchronous elements.

Reflections on a decade of Data Science 101: key insights and lessons learned

Author: Todd Iverson

On the heels of the 10th anniversary of the undergraduate Data Science program at Winona State University, this talk explores the evolution of our introductory course. The course has expanded into a comprehensive survey of the data science field, covering key topics such as data management, visualization, and statistical learning. At Winona State, a non-selective institution, the curriculum has been intentionally designed to accommodate students with varying technical backgrounds. To achieve this, we have integrated low-code tools while providing a gradual introduction to programming. Additionally, this session will examine how the introductory course serves the broader Data Science program, providing a solid foundation for students' future success.

2.7.2 DATA FOR TEACHING

How to settle statistical arguments

Author: Paul Fannon and Vicki Hodgson

When a student thinks you are wrong, we can often win arguments by appealing to authority, but what happens when there is a disagreement with a colleague? The epistimology of statistics, certainly amongst statistical users, is very different to other areas of mathematics. There are many "zombie facts" which are prevalent across sources and which people believe because they have been taught them by authorities. Unfortunately, we often find that users of statistics do not know what is going on "inside the box" so do not have the tools to check the veracity of these facts. I have found that where formal proofs often fail, simulations can be extremely convincing. In this talk I will go through some of the misconceptions I have observed (severall of them my own) and how simulations helped to clarify the situation. Warning: If you teach your students this method my experience is that they can and will highlight all your own mistakes!

UKCOTS25 Abstracts

Synthetic data fidelity: how less can be more

Author: Jools Kasmire

Synthetic data is generated, not observed, and thus includes “made up on the spot” data, random numbers, predictions from complex machine learning models, digital twin simulation output and much more. Fidelity captures how “faithful” synthetic data is to real-world counterparts, often seen as an important (if not the most important) feature of synthetic data. Yet, fidelity is not binary; synthetic data can be faithful in some ways and unfaithful in others, with its usefulness determined by context. For example, synthetic data intended to fix gaps or biases in real-world data must be deliberately unfaithful in some ways. Further, some synthetic data doesn’t try to mimic, replicate or augment real-world data and may not even use real-world data for its generation. Thus, (high) fidelity is not always important. This talk introduces synthetic data and explores the role of fidelity. Focussing on the use of synthetic data for teaching, the talk explains common generation methods and dives into the pros and cons at varying levels of fidelity. The talk concludes by showing how synthetic data provides a valuable alternative to real-world data in educational settings and other situations where reproducibility and transparency must be balanced against privacy protection.

3 Poster Abstracts

3.1 Posters: all breaks with particular focus on Wednesday 25th June, 10:15 – 10:45

Another switch to R: old and new lessons

Author: Anna Riach

Plenty of universities have switched or are considering making the switch to R. We are two years into our change from teaching our undergraduates R instead of SPSS in the School of Biology at the University of Leeds. In making this switch there were lessons I successfully applied that were gathered from speakers at conferences, fellow statistics lecturers at other institutions and papers on the subject, that I will share. I also implemented some new strategies that proved beneficial and naturally had to solve some unexpected problems along the way. One of the reasons for our relatively smooth transition to R is because I have listened and talked to people at conferences like this. I hope this poster serves as a prompt to continue such conversations and make the switch easier for others.

Developing modern undergraduate programmes in Statistics at the University of Glasgow

Authors: Mitchum Bock, Craig Alexander, Eilidh Jack, Jethro Browell

The University of Glasgow has offered undergraduate degrees in Statistics since 1966 and periodically reviews the curriculum of single and joint subject statistics programmes. The last review of undergraduate programmes concluded in 2012 with the creation of the curriculum as it stands today, with individual courses maintained and updated in isolation since then. In 2024, a committee of staff was formed to review honours statistics and has proposed an updated programme with changes in structure and content to reflect modern statistical practice and the skills expected of graduates by employers and required for further study.

This poster presents the proposed updates to the University of Glasgow's BSc and MSci degrees programmes in Statistics and associated joint degrees and invites your feedback as external experts in the teaching of statistics in higher education. The proposal includes changing the names of these programmes to "Statistics & Data Analytics" to reflect their new content.

Proposed curriculum changes include expanding topics in machine learning from a statistical perspective, teaching Python alongside R, and re-focusing practical courses on graduate attributes, such as handling messy/missing/erroneous/"real" data and the ethics and usage of Generative AI. Increasing content in some areas necessitates removing content in others, unless additional teaching resources become available. 30-credits of new courses are proposed, and we invite feedback on what optional content to prioritise retaining in the programme overall.

Structural changes are proposed to improve student experience and reduce assessment burden by merging six 10-credit courses into three 20-credit and examining some courses in the December exam diet. This will reduce the total number of exams sat by students, and distribute them across two diets each year, where currently students sit high number of exams, relative to other University of Glasgow degrees, in a single diet each year. Continuous assessment and project work will continue to be a feature of the programme.

Examining students' perceptions of online learning as compared to face-to-face lectures

Author: Emma Howard

The initial CoVid-19 lockdown in March 2020 presented a unique educational research opportunity whereby undergraduate students took up to six modules in a face-to-face format (pre-lockdown) and online format (during lockdown). In May 2020, 156 students studying mathematics, statistics and applied mathematics modules completed a survey whereby they were asked to categorise each of their module's face-to-face and online formats using pre-defined categories, and rate, on a scale of 1-5, how beneficial they perceived each type of lecture format to be for their learning. This poster presents the results of this study (Meehan & Howard, 2024). Findings indicate that pre-lockdown almost 70% of the lecture formats classified by students were traditional in nature, 20% classified as containing some interactions, and 10% as including group work. Students generally did not perceive great differences in terms of the benefits to learning between the face-to-face and online lecture formats except in the cases where the face-to-face lectures had interactions. This study emphasises that technology is a tool, and the importance of student-to-student and student-to-lecturer interactions. Meehan, M., & Howard, E. (2024). The university mathematics lecture: to record, or not to record, that is the question? Mathematics Education Research Journal, 36, 285-310, https://link.springer.com/article/10.1007/s13394-023-00444-2

Interdisciplinary statistics in HE: deciding what is needed for the discipline

Author: Christoper Aldous Oldnall

Across disciplines, there is a general agreement that a baseline level of statistical knowledge is essential for students to succeed and apply statistical methods effectively. However, what constitutes this baseline remains a subject of debate. Different fields such as mathematics, medical sciences, and the arts and humanities often prioritise different statistical concepts based on their specific applications, resulting in varied expectations for foundational knowledge. This work explores the topics considered "core" across these different academic cohorts, identifying commonalities and divergences in statistical education. A key objective of this research is to develop a structured, adaptable framework that guides the selection of baseline statistical content based on course duration, student pre-existing knowledge, and disciplinary requirements. By designing a "flowchart-style" system, we aim to provide a flexible approach that ensures students receive the most relevant statistical training for their field of study without unnecessary redundancy or gaps. The proposed framework offers a practical tool for educators to tailor statistics curricula to diverse student backgrounds, optimising learning

efficiency while maintaining essential competencies. Ultimately, this work seeks to contribute to a more coherent and adaptive understanding of statistical literacy across disciplines, fostering interdisciplinary dialogue and improving statistical education.

Software week: an approach to developing students' statistical software skills

Author: Liz Grant, Alice Batchelor

The Maths and Statistics Advice (MASA) service at the University of Surrey runs a "Software Week", a week-long series of activities open to all students focused on developing essential skills in R & RStudio, Excel, SPSS and Jamovi. Throughout the week there is a mix of bookable Excel and RStudio workshops, as well as a drop-in "Software Clinic", offering one-to-one support for SPSS and Jamovi. Running the workshops in a single week requires students to focus their learning, and promoting the activities as a week-long event helps highlight to students that statistical software is a key topic that the MASA service can assist students with. Following a successful pilot run last academic year, this year's program has built on participant feedback by:

- Splitting the Excel workshops with each one given a specific focus and more interactive exercises
- Adding an extra R and RStudio workshop for intermediate level users
- Replacing the structured SPSS and Jamovi workshops with more flexible Software Clinics
- Developing additional support resources, hosted on our online learning module

This poster highlights the impact of the program through usage statistics, student feedback, and snapshots of workshop content, showcasing its effectiveness and the improvements made for this academic year. It also provides a practical model for other institutions looking to implement similar initiatives.

Teaching measurement error models to develop statistical thinking

Author: Anarina Murillo

Introductory biostatistics courses are essential for teaching data literacy and statistical thinking. A problem-first approach allows students to draw connections in statistics and data science to their respective fields in medicine or public health. Novel approaches are needed to introduce advanced statistics concepts in these courses. This work explores the effectiveness of teaching measurement error (ME) methods in an introductory biostatistics course to reinforce understanding of bias, uncertainty, and decision-making. A tutorial was developed using the National Health and Nutrition Examination Survey (NHANES) data. We proposed a logistic regression model to predict adult diabetes status based on blood pressure levels adjusted for age, race, calories, alcohol intake, and body mass index (BMI). To imitate real world applications of ME correction, we analyzed a subset of the data (n=600) and simulated random noise to imitate random error in our dataset. Regression calibration (RC) was applied to correct for the simulated error in the model. Students were taught that RC may reduce the effect of ME in statistical models. Students reported improved understanding of statistical

inference. This work demonstrated the utility of an online tutorial on RC to introduce advanced statistical topics in courses addressing bias, uncertainty, and decision-making.

Testing student's ability to be (un)certain

Author: Ulrika Sahlin

Uncertainty quantification can be understood as a pure mathematical procedure and/or a human oriented characterisation of scientific uncertainty given an evidence basis. With the aim to evaluate student's factual, procedural and conceptual knowledge about the latter, we define "uncertaincy" as a person's ability to express her (un)certainty and developed tests of uncertaincy combining performance scores of expert judgement and the ability to justify judgement. Uncertaincy was tested on a course in mathematical and statistical uncertainty analysis for risk engineering students at Lund University, Sweden. At the beginning of the course, the students were tested for their numeracy, including risk literacy and statistical literacy. Each week of the course, the students performed tests for numeracy and for uncertaincy. Uncertainty was evaluated with respect to different objects for uncertainty: facts (categorical variables that, at least theoretically, are directly verifiable), numbers (non-variable quantities) and scientific hypothesis (theories about how the world works). We ask if uncertaincy can be explained by numeracy, if numeracy and uncertaincy were improved during the course, and if there was a difference between the three objects for uncertainty.
