- **Which Stats Package**: SAS, SPSS, R, Statistica or spreadsheets?
- **SPSS for Mixed Models**: Step-by-step instructions.
- **SAS (and R) for Mixed Models**: Free software and resources.
- **P Values Down But Not Yet Out**: Critique of a critique.
- **Journal Impact Factors 2016**: the latest Elsevier/Scopus values

Reprint pdf · Reprint docx

Which Stats Package?

Will G Hopkins, Institute of Sport Exercise and Active Living, Victoria University, Melbourne, Australia. **Email**. Reviewer: Alan M Batterham, School of Health and Social Care, University of Teesside, Middlesbrough, UK. Sportscience 20, i-ii, 2016 (sportsci.org/2016/inbrief.htm#which). Published November 2016. ©2016. Reviewer's Comment.

Updated March 2017. The spreadsheets for controlled trials and crossovers now allow for two covariates, so they are preferable to any statistics package for most interventions and time series.

A browser-based version of the Statistical Analysis System (SAS), *SAS Studio*, is now available as a free "University Edition". It has a point-and-click interface and cool graphics, and it allows easy access to the code or script for development of sophisticated analyses. The only thing it doesn't offer is neural-net modeling, which is available in the full SAS package at extra cost as the *Enterprise Miner*. The only other stats packages in contention that I have used are SPSS, R Studio and Statistica. Each has major limitations for mixed modeling, which is the way you should do all your analyses from now on. In mixed modeling you estimate means via the usual fixed effects, but you also estimate standard deviations via random effects, which can specify individual differences or responses and allow for different errors at different time points with repeated measurement or in different groups or groupings, thereby properly accounting for non-uniformity. This in-brief item summarizes the functionality of SAS Studio, SPSS, R, Statistica, and my spreadsheets for mixed modeling.

SAS Studio has the full suite of parametric modeling procedures, including Proc Mixed and Proc Glimmix. You use Proc Mixed for general linear mixed modeling of the usual continuous dependent variables. Proc Glimmix is for genera*lized* linear mixed modeling, which you need for the more difficult dependent variables: binaries representing classifications or events, counts of anything, and proportions of

anything. Proc Glimmix works just like Proc Mixed; it was introduced a few years ago to improve on the entry-level generalized procedure, Proc Genmod, which has limited random effects and less intuitive output. When you use the point-and-click programming for generalized linear modeling in SAS Studio, only Proc Genmod is invoked, so you have to learn how to write the code with Glimmix. A package of materials updated in this issue introduces you to SAS Studio, SAS coding, and mixed modeling.

In **SPSS** the general linear mixed model does not allow negative variance (negative variance does make sense, especially for individual responses), but otherwise it performs well and its interface is reasonably friendly. SPSS has two generalized linear mixed models: the first is comparable to SAS's Proc Genmod, while the second appears to be an unsuccessful attempt (in SPSS Version 23) at something similar to Proc Glimmix. Resources for mixed modeling in SPSS are available in this issue, but my advice is to learn SAS Studio.

In the free open-source stats package **R Studio** there are two mixed models, lme4 and nlme. Neither offers negative variance, and worse still, they do not provide standard errors for the random effects, so you have no idea of the uncertainty in the standard deviations. Someone developed some code to get the standard errors, but it gives answers different from those in SAS. The other problem is the extreme unfriendliness of the R language, even in the R Studio version. I spent many hours with Alice Sweeting at Victoria University Melbourne trying to figure out how to specify straightforward random-effect models in R. We got some going, but we gave up with multiple

levels of repeated measurement when there was a mix of correlated and uncorrelated random effects. Proc Mixed handled them brilliantly. A summary of some simple mixed modeling with R authored by Alice is available here, but it is unlikely to be updated. See also Alice's blog Sport Statistics R Sweet to develop your skills with this package. R has been integrated with some laboratory hardware and software for data acquisition, which is fine, but learn to use SAS Studio for most or all of your subsequent data processing and modeling.

Statistica has a friendly interface, but I have tried the mixed model on several occasions over the years without success. I'd like to hear from anyone who can make it work for straightforward reliability and controlled-trial analyses.

Finally the **spreadsheets** at this site are useful for straightforward designs with a continuous dependent variable: all you have to do is copy in your raw data. The spreadsheets do log transformation, standardization, and when you add a smallest important effect, magnitude-based inferences. The spreadsheets for controlled-trials, crossovers and comparison of group means allow for different standard deviations in two groups and are thereby equivalent to mixed modeling. Unfortunately they are limited to one covariate; for example, you can't adjust simultaneously for baseline and another

subject characteristic in controlled trials and crossovers. On the other hand, they actually show you graphically what it means to adjust for a covariate, and they provide proper estimates and inferences for the magnitude of the effect of the covariate. To do these things with a stats package requires a lot of work, as you will find when working with SAS Studio. The spreadsheet for reliability is also better than mixed modeling for consecutive pairwise analyses, which is what you want when estimating error of measurement for most tests. The validity spreadsheet also does a great job.

In summary, use my spreadsheets for analysis of a continuous dependent variable with no more than one covariate. Use SAS Studio for everything else–it's a game-changer.

Reviewer's Comment. This article introduces SAS Studio software and pitches it against SPSS, R, and Statistica, plus the spreadsheets available at sportsci.org, for conducting linear mixed modelling. The limitations of the other packages lead to the conclusion that SAS Studio is the package of choice. The spreadsheets–which are validated against SAS output–are recommended for simple analysis of continuous outcomes with a single covariate. I have experienced all of the packages compared in this article and I agree with the author's conclusions and recommendations.

SPSS Mixed Models

Will G Hopkins, Institute of Sport Exercise and Active Living, Victoria University, Melbourne, Australia. **Email**. Reviewer: David S Rowlands, Massey University, Wellington, NZ. Sportscience 20, ii-iii, 2016 (sportsci.org/2016/inbrief.htm#SPSS). Published March 2016. ©2016. Reviewer's Comment

Updated 7 June 2016. I have now provided resources for generalized linear mixed modeling, with worked examples of a count as a dependent variable (Poisson regression) and a proportion as a dependent variable (logistic regression), with repeated measurements on subjects. I presented these resources at a workshop at the University of Bath June 1-3. I have not provided a worked example of a binary dependent variable (i.e., two values only), but it's a simple matter to choose that option with the GEE approach in SPSS. There are some earlier resources for a binary dependent without repeated measurement. There are no resources for analysis of counts or proportions without repeated measurement, but again, it's a simple matter to find and use the right program in SPSS. The zip-compressed file now also con-

tains a slideshow about magnitude-based inference, some of which was shown at Bath and at Leeds and Split. It is based on recent publications about inference. See the next In-brief item on p values below.

I have now extensively updated the files on the use of SPSS for mixed modeling and other analyses that were previously available at this site. The occasion for the update is a workshop I presented at Leeds Beckett University in the UK. Download the Zip-compressed file, in which there is a brief slideshow explaining mixed modeling (edited from previous slideshows at this site), three Word docs with step-by-step instructions and several Excel spreadsheets to import into SPSS when needed. Work your way through the Word docs in this

sequence:

SPSS basics and reliability mixed models.docx

SPSS controlled-trial mixed models.docx

SPSS generalized mixed models.docx

SPSS analysis of binary outcomes.docx (you could skip this one)

In SPSS Version 21 and presumably earlier versions there was a bug in the generalized estimating equations (GEE), such that it gave wrong answers for confidence limits of factor (nominal) fixed effects when a covariate was included in the model. That appears to have been fixed in Version 23 on.

I will present some of the material again at a workshop on mixed modeling at the Vienna meeting of the European College of Sport Science in July. I hope to do something similar, if less extensive, with R for the ECSS workshop, and I will update these files after each workshop.

Reviewer's Comment. This workshop provides a solid grounding in mixed modeling with SPSS for reliability, controlled trials, and generalized mixed models, in clearly described logical annotated workflows. However, I advise you to learn how to use SAS or SAS Studio to take advantage of the better mixed-modeling capacity of SAS.

SAS (and R) for Mixed Models

Will G Hopkins, Institute of Sport Exercise and Active Living, Victoria University, Melbourne, Australia. **Email**. Reviewer: Alan M Batterham, School of Health and Social Care, University of Teesside, Middlesbrough, UK. Sportscience 20, iii, 2016 (sportsci.org/2016/inbrief.htm#SAS). Published June2016. ©2016. Reviewer's Comment

Download the 11.9 MB zip-compressed file of workshop materials on the use of SAS Studio/University Edition and mixed-modeling. The file contains a slideshow introducing mixed modeling (download it separately here), a slideshow about magnitude-based inference (but see the article and slideshow on magnitude-based decisions in the 2020 issue), and of course, the resources on SAS. Put the zipped file where you want the package to reside, right-click on the file and select Extract All. If you use Internet Explorer and get "compressed folder is invalid" when you try to open it, copy the above link into Chrome or Firefox browsers.

The updated slideshow on analysis of repeated measures first presented at the ACSM conference in 2003 is no longer available in the right-hand navigation frame. Access it via this link to a Newstats page.

Updated 20 March 2023. As noted in this In-brief item in 2022, the above zip-compressed file has been updated to make the instructions for getting started consistent with the use of SAS Studio ODA (OnDemand for Academics). The included slideshow on magnitude-based inference has been replaced with a slideshow on sampling uncertainty, also available as a link in this article.

Updated 19 May 2018. Minor updates to the slideshow on mixed modeling. Updates to Sharing folders in SAS Studio.docx in the Getting Started folder to account for changes in the VirtualBox interface.

Updated 1 May 2017. Crossover programs added to the Controlled-trial models folder. These programs reproduce the analyses of the post-only crossover spreadsheet with two covariates.

Updated 6 January 2017. Minor changes to the two slideshows on magnitude-based inference and mixed modeling. Extensive update to the reliability mixed models to include more instructions on graphing with proc sgplot.

Updated 7 December 16. Minor cosmetic changes to some files, especially the spreadsheet to process Poisson and logistic repeated measures.

Updated 4 November 2016. The zip-compressed file now contains improved instructions on installing SAS Studio, instructions on accessing folders and files on your computer from within SAS Studio, a browser page to access help at the SAS site, and a new set of instructions to get started with SAS Studio by analyzing simple statistics for subject characteristics. I have also updated the existing suite of mixed-model analyses, including a major update of the generalized mixed models to include log-hazards and logistic regression for binary variables and magnitude thresholds defined by standardization. A partial summary of corresponding programs for use with the R package authored by Alice Sweeting at Victoria University Melbourne is also available here, but it is unlikely to be updated. See also Alice's blog Sport Sta-

tistics R Sweet to develop your skills with R.

In preparing for the workshop on mixed modeling at this year's ECSS conference, I discovered that the Statistical Analysis System is available in a free version (SAS University Edition) of the version of SAS that runs in a browser window (SAS Studio). This point-and-click version of SAS is much better than the SPSS and R packages. I will therefore make SAS the focus of the workshop, although I will also introduce attendees to the package of SPSS materials already available here (see above) and provide R script for some of the examples. I will also provide materials for SAS to download here before the workshop [now available at the link to the zip-compressed file at the start of this item]. So stay tuned, but meantime follow the instructions in this PDF [now redundant] on how to download and install SAS University Edition. If you are attending the workshop, work your way through the three short videos mentioned in the PDF, and bring your laptop to the workshop with SAS installed and running.

Reviewer's Comment. This update represents very valuable resource. Working through the workshop provides a solid grounding in mixed modelling with SAS Studio software. The examples move beyond linear mixed modelling of continuous outcomes to generalized linear mixed models with a variety of distributions and link functions to analyze binary, count, and proportion outcomes.

P Value Down But Not Yet Out

Alan M Batterham, Will G Hopkins. Health and Social Care Institute, Teesside University, Middlesbrough, UK; Institute of Sport Exercise and Active Living, Victoria University, Melbourne, Australia. **Email**. Reviewer: David S Rowlands, Massey University, Wellington, NZ. Sportscience 20, iv-v, 2016 (sportsci.org/2016/inbrief. htm#Pout). Published March 2016. ©2016. Reviewer's Comment

Updated 1 May 2017. A slideshow explaining p values, magnitude-based inference, and the ASA's policy statement is now available. See the In-brief item in the 2017 issue.

On March 7 the American Statistical Association (ASA) published their eagerly anticipated policy statement on the "context, process and purpose" of p values. The ASA assembled a group of 20 experts for a two-day meeting, facilitated by Regina Nuzzo, author of an influential article published in Nature in 2014 on problems with p values. The meeting was preceded by many months of discussion between group members, and there were multiple iterations of the draft statement following the meeting to arrive at a consensus. Revealingly, the ASA statement notes that "the statement development process was lengthier and more controversial than anticipated". Reading between the lines and examining the social media buzz, it seems to have been quite a bun-fight.

The ASA statement includes six principles...

1. P values can indicate how incompatible the data are with a specified statistical model.
2. P values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
3. Scientific conclusions and business or policy decisions should not be based only on whether a p value passes a specific threshold.
4. Proper inference requires full reporting and transparency.
5. A p value, or statistical significance, does not measure the size of an effect or the importance of a result.
6. By itself, a p value does not provide a good measure of evidence regarding a model or hypothesis.

We welcome the ASA statement as a step forward, but in our view it doesn't go far enough. The p value has utility, according to Principle 1. The use of the word *only* in Principle 3 implies that the p value still has to fall below an unspecified threshold before a conclusion or decision can be reached, but that something else is needed. What exactly? In Principles 5 and 6, it appears that the magnitude of the effect should be taken into consideration. We interpret Principles 1, 3, 5 and 6 together as a requirement for statistical significance (at whatever threshold) before you can assess the magnitude of an effect. We have called such an approach the *conservative null-hypothesis significance test* (NHST) in our recent article "Error rates, decision rates and publication bias with several inferential methods", available online in *Sports Medicine*. In our simulations of controlled trials, conservative NHST has unacceptably high false-positive and false-negative error rates for marginally trivial-small effects

with large sample sizes, along with unacceptably low decision rates and substantial publication bias with small sample sizes. It fares little better than *conventional* NHST, according to which significant effects are substantial and non-significant effects are trivial or even null. The various forms of the magnitude-based approach to inference that we have introduced outperform these two versions of NHST. For an overview, read the abstract of our article in Sports Medicine and an In-brief item in last year's issue of Sportscience.

Invited commentaries by group members have been published alongside the policy statement (available via the supplemental link on the policy statement page). At least two members alluded to problems in making decisions about outcomes. John Ioannidis noted that decisions about an effect based on some "magic threshold" may be biased, regardless of how the threshold is defined. Ken Rothman was opposed to making any decision about an effect, because the decision can end up divorced from the richer information contained in the confidence interval. In our view, decisions about effects are unavoidable: the editor of a journal has to decide whether to publish a study, and the practitioner has to decide whether to implement an intervention. Both decisions involve acceptable uncertainty in the effect, and magnitude-based inference provides a rubric for such decisions. We have shown that the associated bias is negligible, the associated error rates are acceptable, and the decision/publication rates are higher than with NHST.

The authors of the following two excerpts come closest to advocating magnitude-based inference, and like us, they call for abolition of NHST:

> "I have to teach hypothesis testing, since it is so prevalent in biomedical research, but life would be much easier if we could just focus on estimates with their associated uncertainty. The basic artifice of hypothesis testing as a concept is perhaps the root cause of the problem, and I doubt that it will be solved by judicious and

subtle statements like this one from the ASA Board." (Roderick Little)

> "We can and should advise today's students of statistics that they should avoid statistical significance testing and embrace estimation instead… Real change will take the concerted effort of experts to enlighten working scientists, journalists, editors and the public at large that statistical significance has been a harmful concept, and that estimation of meaningful effect measures is a much more fruitful research aim than the testing of null hypotheses. This statement of the ASA does not go nearly far enough toward that end, but it is a welcome start and a hopeful sign." (Ken Rothman)

Reviewer's Comment. This In-brief article provides a commentary, albeit with a tone of consternation, of what should have been a landmark position statement from the American Statistical Association (ASA) supporting the *dis*-use of p values. Alan and Will have summarized the key statements from the ASA and point out the implied caveats, which are effectively escape routes to justify continued *mis*-use of p values. The fallacies and pitfalls of null-hypothesis significance testing have been widely published for over 70 years, yet on a weekly basis I see problems arising from p values. Recently, I struck scientists in a multinational company who made decisions involving millions of dollars based on an inferential decision generated from non-significance with a null-hypothesis test (p value), which later turned out to be misleading. A different decision would have prevailed with magnitude-based inference.

I urge researchers, journal editors and reviewers to carefully read the articles on magnitude-based inference and come to terms with this approach to statistical and inference. You should then share in the responsibility to inform and educate government policy boards, industry, funders, other scientists, and the die-hard statisticians.

Updated 14 Dec 2016. The Journal Metrics link now takes you to a new page, where you can enter individual journal names to get citation statistics or download a workbook of statistics for all journals for the last few years. Elsevier previously referred to the traditional impact factor as *impact per paper* (IPP); now they are calling it the *CiteScore*. My workbook shows the latest (2015) impact factors, along with those for 2014 and 2013, for all the journals in the disciplines of exercise and sport science and medicine, sorted by title and by impact factor. This year's scores appear to have been updated from those published earlier this year, judging by the fact that most scores are up on last years', but a few high-scoring journals have dropped by ~10%.

As noted in last year's article on journal impact factors, I have abandoned Thomson-Reuters' impact factors in favor of Elsevier's, which are derived from a bibliographic database (Scopus) more relevant to sport and exercise science, and which are freely available in a very large workbook (22 MB) at Journal Metrics. I have extracted the values for our journals into a user-friendly and much smaller workbook (23 KB), which has spreadsheets sorted by journal title and by 2015 impact factor. I don't intend write a full article on the impact factors from now on.

There is a problem with this year's impact factors, which are almost all down on last year's. As in previous years, the values were compiled in June of the current year (2016) for articles published in journals to the end of the previous year (2015), so the Scopus database can't be fully up to date yet. I expect there will be an update, and if there is, I will update the spreadsheets.

———