requirements and speed up training, data is input to the model in batches (or subgroups of observations) with an update of model parameters after each batch. One epoch of training is complete when the model has seen all batches that make up the dataset.

```
history <- model %>% fit(x_train, y_train,
      epochs = 200, batch_size = 32,
          validation_data = list(x_val,y_val))
```

In practice, it will be important to repeat the analysis over a broad range of choices of the various hyperparameters (eg. number of layers, number of nodes per layer, regularization method, etc.) and choose the hyperparameters for use in the final model that produce the best results.

Evaluation and prediction

To test the generalizability of the ultimately chosen model, the trained model is evaluated on data not seen by the model during the training process. The root mean squared error of a regression estimate or the accuracy of classifications can be determined, providing a measure of how effective the model is at making predictions on previously unseen data.

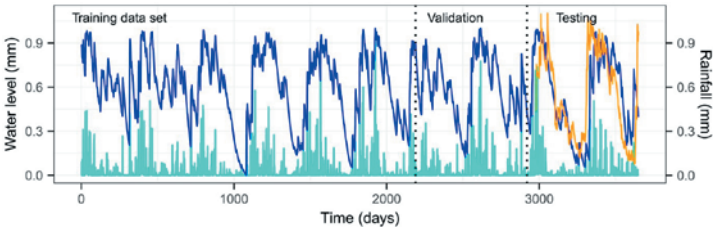
```
model %>% evaluate(x_test, y_test)
```

Finally, predictions can be made with the trained model. A list of predicted y-values or classes for each item of x_test can be found with:

```
model %>% predict(x_test)
      model %>% predict_classes(x_test)
```

Application

We recently used Keras to model time series corresponding to daily measurements of groundwater levels as a function of rain and other climate variables. As a simple example, the following figure shows the observed groundwater levels in blue, along with rainfall measurements in turquoise. The split of the data into training, validation and testing segments can be seen. The orange line shown in the testing segment represents the predictions constructed by an LSTM using the training and validation components of the dataset. See Clark et al (2020) for more discussion.



References and additional resources

All code used in this article can be found here.

Chollet, F. & Allaire, J. J. (2017). *R interface to keras*. https://github.com/rstudio/keras

Chollet, F. & Allaire, J. J. (2018). *Deep learning With R*. Manning Publications Company.

Clark, S., Hyndman, R.J., Pagendam, D. and Ryan, L.M. (2020). Modern strategies for time series regression. International Statistical Review, 88, pp.S179-S204.

Ghatak, A. (2019). *Deep Learning with R*. Springer.

Goodfellow, I., Bengio, Y. & Courville, A. (2016). *Deep Learning*. MIT Press. http://www.deeplearningbook.org

The following webpages provide a wealth of explanations, guidance and examples of deep learning in R:

Keras in R: https://keras.rstudio.com/index.html

Examples of applications: https://blogs.rstudio.com/ai/gallery.html

More basic guidance: https://tensorflow.rstudio.com/guide

STRengthening Analytical Thinking for Observational Studies (STRATOS): Report from the Causal Inference Topic Group (TG7)

Ingeborg Waernbaum1, Bianca De Stavola2, Vanessa Didelez3, Erica Moodie4, Saskia le

Cessie5,6, Els Goetghebeur6 on behalf of STRATOS TG7

1 Department of Statistics, Uppsala University, SW.
2 GOS Institute of Child Health, University College London, U.K.
3 Department of Mathematics and Computer Science, University of Bremen and Department of Biometry and Computer Science, Leibniz Institute for Prevention Research and Epidemiology, Germany.
4 Department of Epidemiology, Biostatistics and Occupational Health, McGill University, CA.
5 Department of Clinical Epidemiology and Department of Biomedical Data Sciences, Leiden University, NL.
6 Department of Applied Mathematics, Computer Science and Statistics, Ghent University, BE.

We are pleased to provide a brief update on activities of the causal inference topic group, TG7 of the STRATOS initiative with the objective to support practicing statisticians with guidance on implementing novel statistical methods in the expanding research area of causal analysis. Statistical analysis targeting causal estimands have been under intense development in the last decades, see e.g. the book by Hernan and Robins for an accessible introduction [1]. Performing accurate analysis with justified underlying assumptions remains challenging, however, and revealed itself over again over the course of the COVID-19 pandemic.

In "Formulating causal questions and principled statistical answers" [2], TG7 provides a roadmap for empirical scientists to

perform statistical analysis for causal inquiries. The introductory paper targets causal estimands for specific scientific inquiry and shows how various estimation approaches compare in principle and in an applied setting. Simulating the counterfactual worlds yields gold standard answers which help see the consequence of various analysis choices. The field gains new momentum as its importance beyond observational studies, gets emphasized in the regulated clinical trials world. Citing from addendum ICH E9 (R1) on estimands and sensitivity analysis in clinical trials from the European Medical Agency

"An estimand is a precise description of the treatment effect reflecting the clinical question posed by a given clinical trial objective. It summarizes at a population level what the outcomes would be in the same patients under different treatment conditions being compared."

The tutorial paper [2] is accompanied by ready to go teaching material for a companion course accessible at the group's website www.ofcaus.org. Exercises cover both introductory discussion topics and data analysis involving outcome regression, matching, propensity score regression and weighting, plus combinations thereof besides instrumental variable analyses. Suggested solutions of practicals include R, Stata and SAS.

One current project develops a tutorial for causal survival analysis along similar lines, with due attention for covariate dependent censoring of the time to event outcome. We zoom in on the difference between potential survival curves as the estimand of interest. The role of inverse probability of treatment and/or censoring, outcome regression model (mis-)specification are studied and compared. An introduction to the use of potential survival times is provided in the tutorial paper [3] of the survival analysis topic group, TG8, see also [4] for an early use of potential survival time, and the contribution by Vanessa Didelez in causal mediation analysis for survival outcomes [5].

Led by Saskia le Cessie and Els Goetgehbeur from TG7, the STRATOS initiative has become an active partner in the SISAQOL-IMI (Setting International Standards of Patient-Reported Outcomes and Quality of Life Endpoints in Cancer Clinical Trials – IMI; https://www.sisaqol-imi.org/) consortium. The consortium aims to generate recommendations to standardize the use, analysis, and interpretation of patient reported outcome (PRO) data in single arm studies as well as randomized cancer trials. For this purpose, we collaborate with academic institutions as well pharmaceutical companies, FDA and EMA staff as well as patient and clinician representatives led by an EORTC team.

The members of TG7 have participated in Stratos-organized sessions at various conferences. A selection of presentations: Erica Moodie, Foundations in Causal Thinking for Health

Data Statisticians at the Italian Region Biometric Conference; Vanessa Didelez presenting Causal Inference for Survival Outcomes: a Censored Edition, at the Society for Epidemiologic Research – SER, Bianca De Stavola, Saskia le Cessie and Ingeborg Waernbaum giving invited presentations: Tutorial in Causal Inference I-III at the RSS 2020.

We are proud of Bianca de Stavola, who received the Bradford Hill Medal at the RSS 2021 conference with the motivation: "… for her innovative development of methods for analyzing longitudinal epidemiological datasets and for her work in applying causal analysis in medical statistics".

From here we look forward to new exciting projects, meetings and collaboration!

References

[1] Hernan, M. A., and J. Robins. "Causal Inference: What if. Boca Raton: Chapman & Hill/CRC." (2020).

[2] Goetghebeur, Els, et al. "Formulating causal questions and principled statistical answers." Statistics in Medicine 39.30 (2020): 4922-4948.

[3] Kragh Andersen, Per, et al. (2021). Analysis of time-to-event for observational studies: Guidance to the use of intensity models. Statistics in Medicine 40.1 (2021): 185-211.

[4] Chen, Pei-Yun, and Anastasios A. Tsiatis. "Causal inference on the difference of the restricted mean lifetime between two groups." Biometrics 57.4 (2001): 1030-1038.

[5] Didelez, Vanessa. (2019). Defining causal mediation with a longitudinal mediator and a survival outcome. Lifetime data analysis, 25(4), 593-610.

Region News

Australasian Region (AR)

Annual General Meeting and Mini Conference

The IBS-AR annual general meeting was held in December with 30 members joining via Zoom. Our president Vanessa Cave provided an update and we elected James Curran as President elect and Garth Tarr as Bulletin Correspondent. Congratulations to James and many thanks to Rina Hannaford for her service to the society as she steps down from the Bulletin Correspondent role. We continue to be very ably represented with the re-election of Ian Renner (Treasurer), David Baird (Secretary), Sam Rogers (Membership secretary), Emi Tanaka (Social Media Coordinator) and Hans Hockey (Website manager).

Following the AGM we took the opportunity to have a online mini conference, with a diverse selection of interesting invited talks:

- Anjali Gupta - Variation in hyperspectral imaging data for inks

- Matthew Schofield - Estimating abundance with capture-recapture: the importance of model, estimator, and prior choice

- Ruth Butler - Archaeologist, Detective, Counselor... The many roles of a Biometrical Consultant

- Ellis Patrick - Extracting meaningful information about disease from state-of-the-art cell imaging assays

Joint IBS-AR/SEEM Regional Conference, Bay of Islands, NZ

27 November – 1 December 2023

The Joint IBS-AR/SEEM Regional Conference was to be held at the end of 2021, however given the global uncertainties around COVID-19 it was postponed for two years to November 27th to 1st December 2023 at the Copthorne Hotel and Resort, Bay of Islands, Waitangi, New Zealand. It will be a joint conference with the Statistics in Ecology and Environmental Monitoring (SEEM) meeting. Please make a note of this in your diaries!

As in the past, the conference will look at statistics research and applications in medicine and public health, agriculture and environment, genetics, natural sciences and education.

Brazilian Region (RBras)

Annual Conference

The 65th RBras Annual Conference was held as an online event on September 10-11, 2021 (https://rbras.org.br/65a-rbras-e-19o-seagro/). It was organized by the Department of Exact Sciences, Luiz de Queiroz College of Agriculture (ESALQ), University of São Paulo, with the collaboration of several RBras members and had national and international contributions. We would like to thank all the attendees, speakers, scientific committee and organizers.

The next annual meeting of RBras will be held in Florianópolis, Santa Catarina, Brazil, from November 16-18, 2022. We are grateful to all committee members from the Federal University of Santa Catarina (UFSC) for their willingness and help with the organization. All the IBS community is welcome to attend this conference in the same place where the IBS 2010 was organized in a very successful way.

RBras and the Brazilian Journal of Biometrics

The RBras is pleased to announce the Brazilian Journal of Biometrics (https://biometria.ufla.br/index.php/BBJ), a journal that in 2022 completes 40 years, has become the official journal RBras in July 2021. The journal is open access and does not charge publication fees. Its publisher continues to be the Federal University of Lavras (UFLA), Minas Gerais, Brazil. The Editors would like to invite all of you to consider submitting a paper to our journal.

RBras and the Ômega Talks

Since September 2021, every last Wednesday of each month the RBras has a space in the Omega Talks, where the speakers are invited by the RBras. The Omega Talks are organized by Dr. Paulo Justiniano Ribeiro Junior, Dr. Wagner Bonat e Dr. Walmes Zeviani from Federal University of Paraná (UFPR), RBras members. All the videos can be found In the RBras YouTube page: https://www.youtube.com/c/RBras-IBS/videos.