Practicum - Consultation: 05.05.2022, 09-11

Please upload all files to OLAT on time.

What	File name	When
Individual project	"05worksheet-Your-Name.zip"	09.05.2022 at 7 am.
Group solutions	"05worksheet-Group-Name.zip"	09.05.2022 at 7 am.
Group contribution	"06contribution-Group-Name.zip"	10.05.2022 at 22 pm.

Individual tasks
Your 05worksheet-Your-Name.zip file contains reproducible code necessary to generate your results for both 5A and 5B parts and your report together with the resulting pdf-file.

Exercise 1 (Individual project (Part 5A))

Elicitation of the prior distribution for placebo is described in Baeten et al. (2013) [(appendix)] in the following way:
"From a review of antitumor necrosis factor (TNF)-α treatment in ankylosing spondylitis, historical data were available from eight randomized placebo controlled clinical trials in ankylosing spondylitis patients. The earliest timepoint assessed in this review was 12 weeks after dosing. Assuming a stable placebo response rate between weeks 6 and 12, these data were used in the derivation of the historical data prior.
A random effects meta-analysis of the 8 historical trials was performed assuming exchangeable placebo response rates on the logit scale. Using this model, the predictive distribution for the proportion of responders on placebo in a new study was derived, leading to an estimated response rate of 25% (and a 95% credible interval of 13% to 40%). For ease of use and interpretation, this predictive distribution was approximated by a Beta density with matching mean and standard deviation."
Assume that the data for 8 historical studies subject to placebo are as follows:
Total number of observations n_i in each study subject to placebo $i = 1, \ldots, 8$:
`pl_total←(107,44,51,39,139,20,78,35)`
Number of cases x_i in each study subject to placebo $i = 1, \ldots, 8$:
`pl_case←c(23,12,19,9,39,6,9,10)`

1. Apply the logit-transformation to $p_i = x_i/n_i$ to get an approximately normal distribution of logit-transformed rates (Held and Sabanés Bové, 2020)[p. 64]

2. Implement the code for a random effects meta-analysis in JAGS using the code below

```
[01]    pl1_modelString ← "
```

```
[02]    model{
[03]    for(i in 1:length(y)){
[04]    y[i] ~ dnorm(theta[i], prec_s[i]);
[05]    theta[i] ~ dnorm(mu, prec_tau);
[06]    }
[07]    theta_new ~ dnorm(mu, prec_tau); # predictive distribution for theta

[08]    p_new<-exp(theta_new)/(1+exp(theta_new)); # predictive distribution
at the probability scale
[09]    mu ~ dnorm(0.0, 1.0E-4);
[10]    prec_tau ~ dgamma(1.0E-3, 1.0E-3); # just our assumption
[11]    }
[12]    "
```

3. Explain what is being done in each line of the code.
4. Summarize the posterior predictive distribution for p contained in `p_new`.
5. Does your estimate of the historical response rate for placebo and its 95% credible interval agree well with those reported in Baeten et al. (2013) [(appendix)] "... the predictive distribution for the proportion of responders on placebo in a new study was derived, leading to an estimated response rate of 25% (and a 95% credible interval of 13% to 40%)."?
6. Apply the function for moment matching derived in Exercise 2 of Worksheet 2 to arrive at elicited values for the parameters α and β of a Beta prior.
7. Report your estimates for α and β.
8. Explain the approach to a client.

Exercise 2 (Individual project (Part 5B))

Elicitation of the prior distribution for treatment is described in Baeten et al. (2013) [(Appendix)] in the following way: "The prior distribution for the proportion of responders in the active group was also a Beta distribution. One of the parameters was set to 1. The other parameter was chosen such that there was an approximately 50:50 chance that the responder rate on active treatment would be greater than the responder rate on placebo (based on the prior distributions only). Thus a Beta(shape1=0.5, shape2=1) distribution was chosen."

1. Following the above instructions, develop R code and justify the parameter choice for the treatment group. Recall that for the placebo group a Beta(shape1=11, shape2=32) prior distribution was chosen.
2. Explain what have you done to a client.

Group tasks
Your 05worksheet-Group-Name.zip (one per group) file contains reproducible code necessary to generate your results and your report together with the resulting pdf-file, which can contain scans of your handwritten solutions. List the names of students who contributed to the solution of group tasks.

Exercise 3 (Normal example in JAGS)

Run the code provided in the file `05normal_example_JAGS.R` and explore the interfaces in R to JAGS. Comment your findings.

Exercise 4 (Logistic regression in JAGS)

Extend the code available in the file `05normal_example_JAGS.R` to deal with the logistic regression example for mice data from Collett (2003, p.71) provided in Table 1.

x	y	n
0.0028	26	28
0.0028	9	12
0.0056	21	40
0.0112	9	40
0.0225	6	40
0.0450	1	40

Table 1: Mice data from Collett (2003, p.71).

Compare the output provided by the classic logistic regression and the Bayesian inference. What are the differences?

Exercise 5 (CODA for the logistic regression in JAGS)

Run the logistic regression in JAGS for mice data from Collett (2003, p.71) provided in Table 1, which you developed in Exercise 4 above.

1. Rank plots are histograms of ranked posterior draws (ranked over all chains) plotted separately for each chain, which were recommended by Vehtari, Gelman, Simpson, Carpenter, and Bürkner (2021) instead of traceplots. Implement rank plots in R, apply them and interpret the results.

2. Run the following formal convergence diagnostics and interpret the results:

   a) `heidel.diag` from package `coda` (Robert and Casella (2004, p.509), Cowles and Carlin (1996, Section 2.10));

   b) `raftery.diag` from package `coda` (Robert and Casella (2004, p.500), Cowles and Carlin (1996, Section 2.2));

   c) `geweke.diag` and `geweke.plot` from package `coda` (Robert and Casella (2004, p.508), Cowles and Carlin (1996, Section 2.3));

   d) `gelman.diag` and `gelman.plot` from package `coda` (Robert and Casella (2004, p.497), Cowles and Carlin (1996, Section 2.1), Gelman and Rubin (1992), Brooks and Gelman (1998));

   e) `stable.GR` from package `stableGR` (Vats and Knudson (2021)).

3. Modify the original `n.burnin`, `n.iter` and `n.thin` parameters as suggested by the convergence diagnostics and re-run the MCMC simulation. Do the new results differ much from those in the first run?

Hint: When interpreting the results consult also the description of the methods provided in `coda`.

Exercise 6 (ESS)

Run the code from the previous exercise with mice data with only one chain monitoring *beta* under the following two conditions:

1. After an adaptation phase of 1000 and a burn-in of 4000 draw a sample of 1000 observations in one chain with thinning set to 1.

2. After an adaptation phase of 1000 and a burn-in of 4000 draw a sample of 10000 observations in one chain with thinning set to 10.

Both chains contain 1000 observations. Answer the following question in (a) and check the correctness of your answer in (b) below:

(a) For which of the above conditions the ESS estimates will be larger and why?
(b) To check your answer: Apply both the 05ess.R code and the function `effectiveSize` from the `coda` package. Compare the ESS estimates with those obtained with the `n.eff` function from package `stableGR` (Vats and Knudson, 2021). Please report your findings.

Group contributions

Exercise 7 (Group contributions for the lecture on 12.05.2022)

Please prepare a group contribution, which your group will present (ca. 5 min) during the next lecture.

(6.1) History and purpose of meta-analysis

(6.2) Short summary: classical meta-analysis for a binary outcome

(6.3) Short summary: classical meta-analysis for a continuous outcome

(6.4) The use of meta-analysis with time (Cochrane, number of citations)

Make sure that the file 06contribution-Group-Name.zip (one per group) contains the pdf-file and the `R` code you want to present.

Optional exercises just for fun!

Exercise 8 (Examples in JAGS (optional))

JAGS provides a translation of OpenBUGS examples from Volumes 1 and 2 in the classic-bugs examples folder online. Choose one example in JAGS you are interested in. Run the code in JAGS and interpret the results.
Compare the code of the OpenBUGS-examples and its translation to JAGS. Report on the major differences in the code.

Exercise 9 (Bad mixing)

Run the code discussed in the blog in JAGS

`https://sourceforge.net/p/mcmc-jags/discussion/610037/thread/679c03bc/?limit=25#ab4d`
and report on peculiarities connected to bad mixing. Apply CODA checks and discuss their findings.

References

Baeten, D., X. Baraliakos, J. Braun, J. Sieper, P. Emery, D. van der Heijde, I. McInnes, J. van Laar, R. Landewé, P. Wordsworth, J. Wollenhaupt, H. Kellner, J. Paramarta, J. Wei, A. Brachat, S. Bek, D. Laurent, Y. Li, Y. Wang, A. Bertolino, S. Gsteiger, A. Wright, and W. Hueber (2013). Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised, double-blind, placebo-controlled trial. *The Lancet 382*, 1705–1713.

Brooks, S. and A. Gelman (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics 7*(4), 434–455.

Collett, D. (2003). *Modelling Binary Data. Second Edition*. Chapman & Hall/CRC.

Cowles, M. and B. Carlin (1996). Markov Chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association 91*(434), 883–904.

Gelman, A. and D. Rubin (1992). Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science 7*(4), 457–511.

Held, L. and D. Sabanés Bové (2020). *Likelihood and Bayesian Inference: With Applications in Biology and Medicine.* Springer.

Robert, C. and G. Casella (2004). *Monte Carlo Statistical Methods, Second Edition.* Springer.

Vats, D. and C. Knudson (2021). Revisiting the Gelman-Rubin diagnostic. *Statistical Science 36*(4), 518–529.

Vehtari, A., A. Gelman, D. Simpson, B. Carpenter, and P. Bürkner (2021). Rank-normalization, folding, and localization: An improved $\widehat{R}$ for assessing convergence of MCMC (with Discussion). *Bayesian Analysis 16*(2), 667–718.