Practicum - Consultation: 31.03.2022, 09-11

Please upload all files to OLAT on time.

What	File name	When
Individual project	"03worksheet-Your-Name.zip"	04.04.2022 at 7 am.
Group solutions	"03worksheet-Group-Name.zip"	04.04.2022 at 7 am.
Group contribution	"04contribution-Group-Name.zip"	05.04.2022 at 22 pm.

Individual tasks
Your 03worksheet-Your-Name.zip file contains reproducible code necessary to generate your results for both 3A and 3B parts and your report together with the resulting pdf-file.

Exercise 1 (Individual project (Part 3A))

In worksheet 2, you showed that Beta(14.5, 10) is the posterior in the secukinumab group with posterior mean response rate 59.2% and Beta(12, 37) is the posterior in the placebo group with posterior mean response rate 24.5%. The predefined criterion for declaring superiority of secukinumab over placebo requires a posterior probability of at least 95% that the ASAS20 response rate for secukinumab patients is higher than that for placebo patients. Provide the estimate of the posterior probability of superiority (PPS) in Table 1 and its Monte Carlo standard error.

1. Draw a MC sample from the posterior distribution in the Secukinumab group and a MC sample from the posterior distribution in the placebo group.
2. Compute a response rate difference (RRD) based on both MC samples and provide a plot of this distribution.
3. Compute the median and the equi-tailed 95%CrI of the distribution of RRD. Interpret this result.
4. Compute the estimate of PPS. Is this estimate larger than the threshold of 95%?
5. Compute the MC standard error (MCse) of PPS (Held and Sabanés Bové (2020) page 258, Section 8.3.1).

Explain your approach to a client and report your results.

Exercise 2 (Individual project (Part 3B))

Table 1: ASAS20 responders at week 6: data provided explicitly and implicitly in Table 2 of Baeten et al. (2013).

Group	n	Responders x (%)	Posterior Response rate	Difference (S-P)	95%CrI	PPS	MCse(PPS)
Secukinumab	23	14 (60.9%)	59.2%	34.7%	11.5–56.4%	99.8%	?
Placebo	6	1 (16.7%)	24.5%				

Extend the response rate difference (RRD) argument outlined in Exercise 1 (Table 1) to response ratio (RR) and odds ratio (OR). Compute the Monte Carlo standard error for the PPS estimate in each case. Compare and report the results.

Group tasks

Your 03worksheet-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.

Exercise 3 (Conjugate Bayes: analytical derivations)

Assume that $y_1, \ldots, y_n$ are realizations (observations) generated by *iid* random variables which follow a N(m, κ^{-1}) distribution. Moreover, assume that the prior of m follows a N(μ, λ^{-1}) distribution, where κ, μ and λ are fixed (known) constants. Derive analytically (with all constants) the following two distributions:

(a) The prior predictive distribution of one future observation y assuming that no observations have been collected yet.

(b) The posterior predictive distribution of one future observation y_{n+1} given that $y_1, \ldots, y_n$ have been observed.

Exercise 4 (Conjugate Bayesian analysis in practice)

Apply analytical formulas derived in Exercise 3 above to the vector of `Height` (cm) measurements 166, 168, 168, 177, 160, 170, 172, 159, 175, 164, 175, 167, 164 of 13 Swiss females. Assume that $y_1, \ldots, y_n$ are observations generated by N(m, κ^{-1}) distribution with $\kappa = 1/900$. Moreover, assume a N(μ, λ^{-1}) prior for m with $\mu = 161$ and $\lambda = 1/70$.

(a) Plot the prior predictive distribution for one observation y and compute its expectation and standard deviation. Estimate $P[y > 200]$ for one future observation of `Height`.

(b) Plot the posterior predictive distribution for one future observation y_{n+1} given that $y_1, \ldots, y_n$ have been observed and compute its expectation and standard deviation. Estimate $P[y_{n+1} > 200 | y_1, \ldots, y_n]$ for one future observation y_{n+1} of `Height`.

(c) Compare the results obtained for predictive distribution with those obtained for the posterior in Exercise 4 of Worksheet 2. Discuss how much posterior, prior predictive, and posterior predictive distributions differ.

Exercise 5 (The change-of-variables formula)

Consider a random variable X, which follows a Gamma(shape= a, rate= b) distribution with density $f(x) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx)$. Frequently, a Gamma prior is assigned to precision $1/\sigma^2$. Apply the change-of-variables formula twice and derive the density of Y and Z random variables, where $Y = 1/X$ and $Z = \sqrt{Y} = \sqrt{1/X}$. The distribution of Y is the Inverse Gamma (IG) prior for the variance σ^2 and Z is the Square root Inverse Gamma (SIG) prior for the standard deviation σ. Plot these three densities for $a = 1.6$ and $b = 0.4$. In addition, plot the densities of Y and Z in the domain range between 0 and 0.5. Interpret the shape of densities of Y and Z close to 0.

Exercise 6 (Monte Carlo: transformations of random variables)

Let the random variable X follow the target Gamma(shape= a, rate= b) distribution with $a = 1.6$ and $b = 0.4$.
(a) Generate a Monte Carlo sample (*i.i.d* realisations of X) of size $M = 1000$ using `rgamma()` function in `R` and by assuming set.seed(44566).
(b) Given the sample in (a) generate a MC sample from the Inverse Gamma distribution $Y = 1/X$.
(c) Given the sample in (a) or (b) generate a MC sample from the Square root Inverse Gamma distribution $Z = \sqrt{Y} = \sqrt{1/X}$.
For each of X, Y, and Z accomplish the following tasks:

1. Plot a traceplot of the MC sample;
2. Plot a histogram of the MC sample with the overlayed true density curve from Exercise 5 above;
3. Summarize the MC sample by computing the sample mean and the sample median.

Question: What is the relation between the sample medians of X, Y, and Z?
Question: Does this relation also apply to sample means of X, Y, and Z? Why not?

Group contributions

Exercise 7 (Group contributions for the lecture on 07.04.2022)

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

(4.1) Markov Chains in a nut shell.

(4.2) History of MCMC: Gibbs sampler.

(4.3) History of MCMC: Metropolis-Hastings.

(4.4) History of MCMC: Hamiltonian (Stan).

Make sure that the file 04contribution-Group-Name.zip (one per group) contains the pdf-file and the R code you want to present.

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.

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