
STK 4021-9021 Autumn 2025

Bayesian Analysis: the Oblig

This is The Oblig, the mandatory assignment, for STK 4021-9021, Bayesian Analysis, Autumn 2025. It is made available at the course website Wednesday October 15, and the submission deadline is Monday October 27, 13:58, *via the Canvas system*. Reports may be written in nynorsk, bokmål, riksmål, English, or Latin, should preferably be text-processed (for instance with TeX or LaTeX), and must be submitted as a single pdf file. Norwegians are kindly reminded of the fact that de med høy sannsynlighet er i stand til å skrive på norsk. The submission must contain your name and the course (i.e. 4021 or 9021).

The Oblig set contains four exercises and comprises five pages (in addition to the present introduction page, ‘page 0’).

It is expected that you give a clear presentation with all necessary explanations, but write concisely (in der Beschränkung zeigt sich erst der Meister; brevity is the soul of wit; краткость – сестра таланта). Remember to include all relevant plots and figures. These should preferably be placed inside the text, close to the relevant subquestion.

For a few of the questions setting up an appropriate computer programme might be part of your solution. The code ought to be handed in along with the rest of the written assignment; you might place the code in an appendix.

Note that you are given only once chance for delivering an acceptable report. All aids, including collaboration, are allowed, but the submission must be written by you and reflect your understanding of the subject. If we doubt that you have understood the content you have handed in, we may request that you give an oral account.

Importantly, Nils asks you to include one extra self-report page, perhaps most naturally placed at the start, where you tell your readers how the work has proceeded, what was complicated or time-consuming (or fun), and what you might have judged to be relatively easy. Tell us also about collaboration, extra sources you may have used, artificial or not.

Application for postponed delivery: If you need to apply for a postponement of the submission deadline due to illness or other reasons, you have to contact the Student Administration at the Department of Mathematics (email: studieinfo@math.uio.no) well before the deadline.

The mandatory assignment in this course must be approved, in the same semester, before you are allowed to take the final examination.

Complete guidelines about delivery of mandatory assignments, along with a ‘log on to Canvas’, can be found here:

www.uio.no/english/studies/admin/compulsory-activities/mn-math-mandatory.html

Enjoy [imperative pluralis].

Nils Lid Hjort

1. The ratio of rates

A WISE MAN PROPORTIONS HIS BELIEF to the evidence (says David Hume). Consider the distribution with density

$$f(y, \theta) = (1/6)\theta^4 y^3 \exp(-\theta y) \quad \text{for } y > 0,$$

with a positive parameter θ .

- (a) Show that this is really a density and that its mean is $4/\theta$. With i.i.d. observations $Y_1, \dots, Y_n$ from this density, write down the likelihood function, and give a formula for the maximum likelihood estimator $\hat{\theta}_{\text{ml}}$. What is its approximately normal distribution?
- (b) With a $\text{gamma}(a, b)$ prior for θ , show that $\theta | \text{data} \sim \text{gamma}(a + 4n, b + n\bar{y})$, with $\bar{y}$ the sample average. Give a normal approximation to this distribution, for n moderate to large. With $\hat{\theta}_B$ the Bayesian posterior mean estimator, verify the Bernstein–von Mises type result that $\sqrt{n}(\hat{\theta}_B - \hat{\theta}_{\text{ml}})$ tends to zero in probability, with increasing data volume.
- (c) For an illustration, suppose measurements of the above type stem from two laboratories, with parameters θ_1 and θ_2 . Assume for illustration that $n = 20$, for both sets of data, with sample means $\bar{y}_1 = 3.33$ and $\bar{y}_2 = 4.44$. For the prior, set $(a, b) = (1.11, 1.11)$. Simulate from the posterior distributions to find 0.05, 0.50, 0.95 quantiles for the ratio $\rho = \theta_2/\theta_1$. You may also attempt to find a representation for ρ in terms of the F distribution.
- (d) Via simulations, provide similar inference summaries for the ratio $\kappa = M/H$, where $M = \frac{1}{2}(\theta_1 + \theta_2)$ and $H = 2/(1/\theta_1 + 1/\theta_2)$ are the usual mean and the harmonic mean: comment on what you find.

2. Three dice: are they equal or are they different?

GOTT WÜRFELT NICHT, but I should be allowed to do so, with my three dice. They are red, blue, green, and I throw them 60 times each, giving $(x, y, z) = (5, 10, 16)$ for the number of ‘six’. Are the underlying probabilities p, q, r equal, or are they not? Below you are free to use the mathematical result that

$$\int_0^1 u^a (1-u)^b du = \frac{a! b!}{(a+b+1)!} \quad \text{for integers } a, b.$$

Furthermore, to compute such numbers, it is advisable to go via logarithms, as with e.g.

`oi = lgamma(a+1) + lgamma(b+1) - lgamma(a+b+2)`

followed by `exp(oi)`.

- (a) Suppose in general terms there are two candidate models M_1 and M_2 for data D . With start probabilities p_1, p_2 for these models, and $\bar{f}_1, \bar{f}_2$ the likelihoods for D under models 1 and 2 conditions, show that

$$\Pr(M_1 | D) = \frac{p_1 \bar{f}_1}{p_1 \bar{f}_1 + p_2 \bar{f}_2}, \quad \Pr(M_2 | D) = \frac{p_2 \bar{f}_2}{p_1 \bar{f}_1 + p_2 \bar{f}_2}.$$

- (b) Suppose next that under model M_1 , data follow the density $f_1(D | \theta_1)$, with prior $\pi_1(\theta_1)$ for that model's parameter, and similarly for M_2 . Show that

$$\bar{f}_1 = \int f_1(D | \theta_1) \pi_1(\theta_1) d\theta_1, \quad \bar{f}_2 = \int f_2(D | \theta_2) \pi_2(\theta_2) d\theta_2.$$

- (c) You should now apply this machinery to my three dice. Explain that the likelihood function is

$$f(x, y, z | p, q, r) = A p^x (1-p)^{n-x} q^y (1-q)^{n-y} r^z (1-r)^{n-z}, \quad \text{with } A = \binom{n}{x} \binom{n}{y} \binom{n}{z}.$$

Model 1 says that $p = q = r$, and Model 2 says that no, the three probabilities are not equal. With the probabilities taken to have uniform priors, find formulae for

$$\bar{f}_1 = \int_0^1 f_1(x, y, z | p) \pi(p) dp, \quad \bar{f}_2 = \int_{[0,1]^3} f_2(x, y, z | p, q, r) \pi(p) \pi(q) \pi(r) dp dq dr.$$

Here f_1 and f_2 as above but under the two model conditions. Compute the two posterior model probabilities, starting with 50-50 chances for the two models.

- (d) Expand your horizon: we now contemplate using also models 3a, $p = q$ but r different; 3b, $p = r$ but q different; 3c, $q = r$ but p different. Explain how the above can be properly generalised. Compute the posterior probabilities for these five models, starting with equal prior chances for them.
- (e) The setting above is clearly in the category of 'toy examples', but the methods and lessons can of course be used for various sets of real data. To illustrate this, consider the data below, counting how many Australians among those having graduated in 1940 were still alive in 1990, sorted in 2×2 categories according to gender and type of education. Treat these as four binomial tables, with underlying probabilities p_1, p_2, p_3, p_4 . Though there are further submodels to consider, concentrate here on Model 1, that $p_1 = p_2 = p_3 = p_4$, and Model 2, that no, these four are different. With uniform priors for the probabilities, compute $\Pr(M_1 | \text{data})$ and $\Pr(M_2 | \text{data})$.

```

11 of 25 :  men, arts
12 of 19 :  men, science
15 of 18 :  women, arts
 6 of  7 :  women, science

```

3. CO2 uptake for Quebec and Minnesota plants

EVEN A PLANT IS A LIVING BEING, THEREFORE we must also consider its mind and soul (says Mwanandeke Kindembo). A study on oxygen uptake for the grass plant *Echinochloa crus-galli* (barnyard grass), a C_4 photosynthetic species, involved Quebec plants (cooler climate) and Mississippi plants (warmer climate), sorted into chilling or non-chilling conditions. With 3 plants for each of the 2×2 categories we have measurements for $n = 12$

plants, summarised in the table below. In the present exercise we do not go into the questions of how the photosynthetic uptake varies with ambient CO₂ concentration, but consider the sample means $\hat{\xi}_i$ and sample standard deviations $\hat{\sigma}_i$, for the twelve plants, averaged over 7 measurements.

Our focus here is on the 12 standard deviations, given in the second column of the table below. From standard theory, assuming normality, the empirical variances $\hat{\sigma}_i^2 = (1/6) \sum_{j=1}^7 (x_{i,j} - \bar{x}_i)^2$ have chi-squared distributions, in the precise sense that $\hat{\sigma}_i^2/\sigma_i^2 \sim \chi_m^2/m$, with $m = 6$. For the analysis below it will be convenient, both for the mathematics and the prior-to-posterior calculations, to transform to $\theta_i = 1/\sigma_i^2$ scale. In other words, starting from observed $Y_i = \hat{\sigma}_i^2 \sim \sigma_i^2 \chi_m^2/m$, we go via $Z_i = m\theta_i Y_i \sim \chi_m^2$, carry out the analysis for the θ_i , and then transform back to the original $\sigma_i = 1/\theta_i^{1/2}$ scale.

xi	sigma	
33.229	8.215	Qn1
35.157	11.004	Qn2
37.614	10.350	Qn3
29.971	8.335	Qc1
32.586	10.321	Qc2
32.700	11.337	Qc3
24.114	6.485	Mn1
27.343	7.653	Mn2
26.400	8.694	Mn3
12.143	2.187	Mc1
17.300	3.049	Mc2
18.000	4.119	Mc3

- (a) When $Z = m\theta Y$ has the χ_m^2 distribution, with its density $g_m(z)$, show that Y has density

$$f(y, \theta) = g_m(m\theta y)m\theta = \frac{1}{2^{m/2}\Gamma(m/2)}(m\theta)^{m/2}y^{m/2-1}\exp(-\tfrac{1}{2}m\theta y) \quad \text{for } y > 0.$$

- (b) If now θ is given the gamma(a, b) prior, with density proportional to $\theta^{a-1}\exp(-b\theta)$, show that $\theta | y \sim \text{gamma}(a + \frac{1}{2}m, b + \frac{1}{2}my)$. Deduce in particular that when $\theta_1, \dots, \theta_n$ are seen as i.i.d. from this gamma(a, b), then the posterior mean estimators are

$$\hat{\theta}_i(a, b) = \frac{a + \frac{1}{2}m}{b + \frac{1}{2}my_i}.$$

Comment on the case where a and b are small.

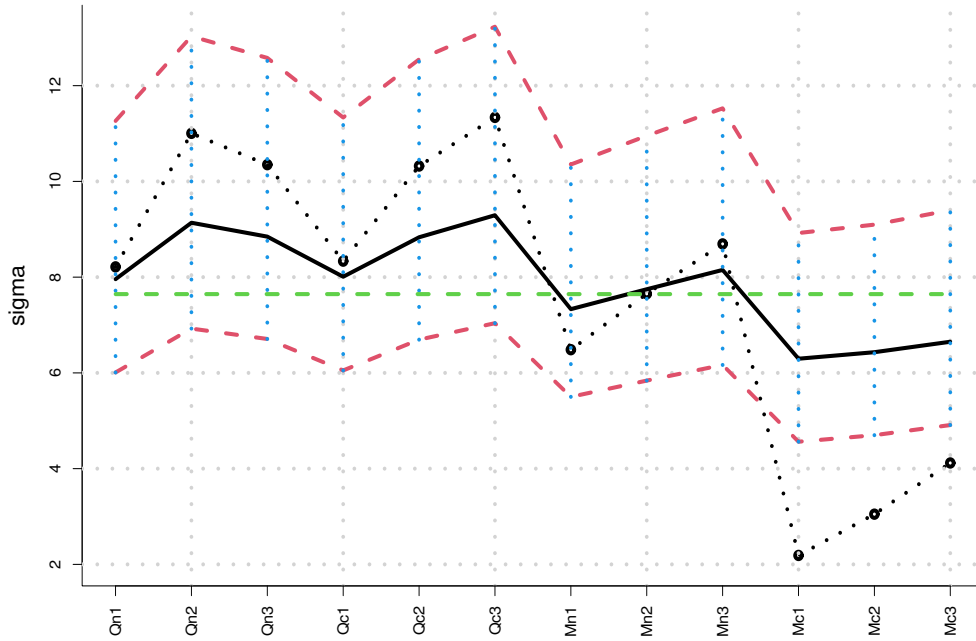
- (c) Since there is no clear recipe for eliciting (a, b) , prior to data, it makes sense to estimate them from the data, leading then to the empirical Bayes estimates

$$\hat{\theta}_{i,\text{eb}} = \hat{\theta}_i(\hat{a}, \hat{b}) = \frac{\hat{a} + \frac{1}{2}m}{\hat{b} + \frac{1}{2}my_i}.$$

Find the marginal density $\bar{f}(y_i, a, b)$, and show that the log-likelihood function may be written

$$\ell(a, b) = \sum_{i=1}^n [a \log b - (a + \frac{1}{2}m) \log(b + \frac{1}{2}my_i) + \log \Gamma(a + \frac{1}{2}m) - \log \Gamma(a)].$$

Maximise this numerically, which involves programming the log-likelihood function (use the `lgamma` function), and then using tools akin to `nlm` in R. I find the maximum likelihood estimates to be $(\hat{a}, \hat{b}) = (6.637, 385.948)$. From this give a plot (i) for $(Y_i, \hat{\theta}_{i,\text{eb}})$, and (ii) mapped back to the sigma scale, for $(\hat{\sigma}_i, \hat{\sigma}_{i,\text{eb}})$. Comment on what you find.



For the twelve plants: the raw data $\hat{\sigma}_i$, ranging from 2.18 to 11.34; the full Bayes medians via the black full line; and Bayesian 90 percent credibility intervals (vertical dotted lines). The overall mean $\bar{\sigma}$ is the horizontal line.

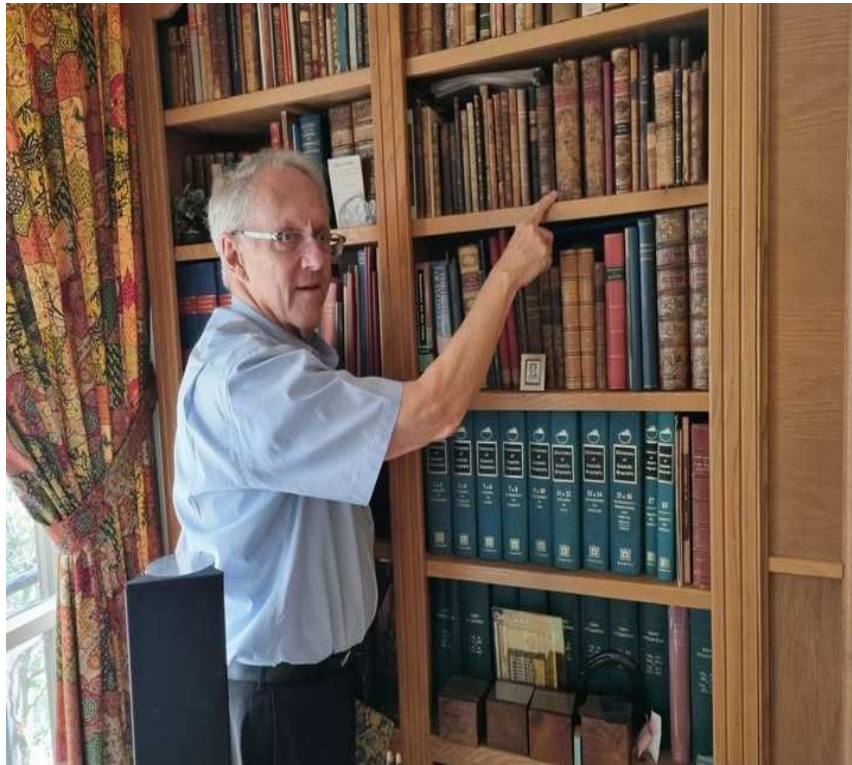
- (d) For a full Bayesian analysis of the twelve empirical standard deviations, let (a, b) have the non-informative prior $(1/a)(1/b)$, and carry out a MCMC, to produce say $\text{sim} = 10^5$ (dependent) realisations (a_j, b_j) from $\pi(a, b | \text{data})$. Give estimates and standard deviations from this output, (i) for a and b ; (ii) for a/b and $a/b^{1/2}$. Comment on how this matches the frequentist results from maximum likelihood theory.
- (e) For each simulated (a_j, b_j) from the posterior, simulate $(\theta_1, \dots, \theta_n)$, and explain that this gives us sim realisations of $\theta_i | \text{data}$, for each of the twelve plants. Use these to produce posterior quantiles, say B_{05}, B_{50}, B_{95} , for 5, 50, 95 percent, for each of the twelve distributions θ_i . Convert these back to the sigma scale, and give a table, perhaps with 3 decimals, of $(\hat{\sigma}_i, \hat{\sigma}_{i,\text{eb}}, q_{05}, q_{50}, q_{95})$, with the 5, 50, 95 percent quantiles on this scale.

- (f) Comment on the not-quite-symmetric form of the intervals, and report briefly on studying histograms for a few of the $\sigma_i = 1/\theta_i^{1/2}$. Check also the degree to which the $\hat{\sigma}_{i,eb}$ are close to the full-Bayes posterior medians q_{50} .
- (g) Construct a version of the figure above, with 90 percent credibility intervals for the twelve σ_i , and the posterior medians placed inside these. Added on are the original twelve $\hat{\sigma}_i$, along with the overall mean of these, plotted as the horizontal line.
- (h) When first looking at the $\hat{\sigma}_i$ of the table, they look somewhat different. But could they perhaps be equal or almost equal, after all?

4. Cooling of newborns

IN THE HJORT–STOLTENBERG BOOK (well, preliminary version of), go to PartTwo and find Story #1, concerning the life-and-death cooling of newborns debate. Carry out as much as you can and deem appropriate, for questions posed there, among the *Bayesian aspects* of that story (i.e. you may bypass the confidence distributions parts).

And, by the way:



*Professor Stephen Stigler, the world leading historian for all matters and themes pertaining to probability theory and statistics, and also owner of the world's most grandiose collection of original books, journals, documents, letters (dating back many centuries), points to his copy of *An Essay Towards Solving a Problem in the Doctrine of Chances*, Thomas Bayes, 1763 (Chicago, October 2025; amateur hjemme-hos photo by Nils).*