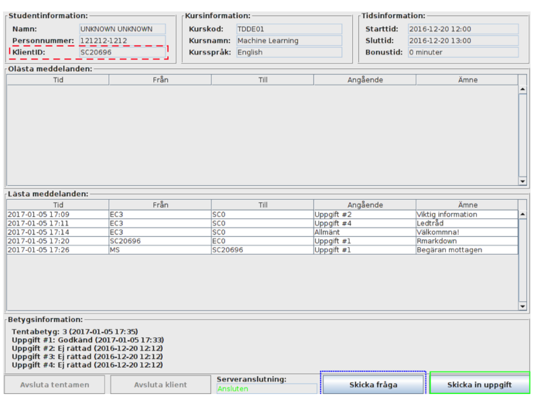
LINKÖPING UNIVERSITY
Dept. of Computer and Information Science
Division of Statistics and Machine Learning
Bertil Wegmann

2022-06-03
Bayesian Learning, 6 hp
732A73/732A91/TDDE07

Computer Exam
Bayesian Learning (732A73/732A91/TDDE07), 6 hp

Time:	8-12
Allowable material:	- The allowed material in the folders given_files in the exam system. - Calculator with erased memory
Teacher:	Bertil Wegmann. Phone: 070 – 1128321 and through the Communication client.
Exam scores:	Maximum number of credits on the exam: 40.
Grades (732A73/732A91):	A: 36 points B: 32 points C: 24 points D: 20 points E: 16 points F: <16 points
Grades (TDDE07):	5: 34 points 4: 26 points 3: 18 points U: <18 points

INSTRUCTIONS:
When asked to give a solution on **Paper**, give that answer on physical papers supplied with the exam.
Each submitted sheet of paper should be marked with your *Client ID* from the *Communication Client*.
The client ID is the code in the `red` dashed rectangle in the figure below.
All other answers should be submitted in a single PDF file using the *Communication Client.*
Include important code needed to grade the exam (at each sub-problem of the PDF).
Submission starts by clicking the button in the `green` solid rectangle in the figure below.
The submitted PDF file should be named *BayesExam.pdf*
Questions can be asked through the Communication client (`blue` dotted rectangle in the figure below).
Full score requires clear and well motivated answers.



1. Demand of a product (*Total credits: 11p*)

Problem (a) should only be solved on **Paper**. A company sells a product. Let Q_i = the demanded quantity of the product in month i, which is assumed to follow a Poisson distribution: $Q_i|\theta \overset{iid}{\sim} \text{Poisson}(\theta)$ and $p(q_i|\theta) = (q_i!)^{-1}\theta^{q_i}e^{-\theta}$. The demanded quantity for five months one year was: $q_1 = 322, q_2 = 248, q_3 = 385, q_4 = 341, q_5 = 310$. A Bayesian analysis resulted in the following gamma posterior distribution for θ with parameters $\alpha_n = 2326$ and $\beta_n = 7$: $p(\theta|q_1,\ldots,q_5) \propto \theta^{\alpha_n-1}e^{-\beta_n\theta}$.

(a) *Credits: 3p.* Derive the prior distribution that was used for θ.

(b) *Credits: 3p.* Simulate 10000 draws from the posterior predictive distribution of the demanded quantity in the next month, Q_6, and plot a histogram of the draws. Compute $\Pr(Q_6 > 350|q_1,...,q_5)$

(c) *Credits: 5p.* The company needs to decide how much of the product to keep in stock for the sale in the next month. The utility function is of the form

$$u(a,Q_6) = \begin{cases} p \cdot Q_6 - (a - Q_6) & \text{if } Q_6 \le a \\ p \cdot a - 0.1\,(Q_6 - a)^2 & \text{if } Q_6 > a \end{cases}$$

where $p = 15$ is the sale price for the product and a (positive integer) is the stock held for the next month. This utility function is given in the file `ExamData.R`. Use simulation to find the optimal a using a Bayesian decision approach.

2. Regression (*Total credits: 12p*) The file `Mollusc.RData`, which is loaded by the code in `ExamData.R`, contains data on 36 different molluscs which live in water. For each mollusc we have observed y_i = *the length in mm for mollusc i*, x_{1i} = *the age in days for mollusc i*, x_{2i} = *the average water temperature where mollusc i has grown in degrees Celsius*. The data matrix contains a column *intercept* with ones to include an intercept in the model. Now, use `BayesLinReg.R` in `ExamData.R` to sample from the joint posterior distribution of the model parameters in the following Gaussian linear regression

$$y = \mu + \varepsilon, \quad \varepsilon \sim N(0,\sigma^2)$$

$$\mu = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_1^2 + \beta_3 \cdot x_2 + \beta_4 \cdot x_2^2 + \beta_5 \cdot x_1 \cdot x_2$$

Simulate 10000 draws from the joint posterior. Use the following prior distributions for $\beta = (\beta_0, \beta_1, \ldots, \beta_5)$ and σ^2:

$$\beta|\sigma^2 \sim N(0, 10^2\sigma^2 I_6)$$
$$\sigma^2 \sim Inv - \chi^2(1, 100^2)$$

(a) *Credits: 2p.* Compute the posterior mean and 99% equal tail credible intervals for all parameters in β and interpret the interval for β_1.

(b) *Credits: 1p.* Compute the posterior mean and posterior median of the standard deviation σ.

(c) *Credits: 3p.* Compute 95% equal tail posterior probability intervals for the expected length μ on a grid of values of x_1 for molluscs with $x_2 = 27$. The grid of values of x_1 shall span between the lowest and highest values of x_1 with the distance 0.1 between any pair of values on the grid. Plot the lower and upper limits of the posterior probability intervals as a function of x_1.

(d) *Credits: 2p.* Investigate if the effect on y from x_1 depends on x_2.

(e) *Credits: 2p.* Compute by simulation the posterior predictive distribution of y for a new mollusc (not in the data) with $x_1 = 50, x_2 = 25$ and plot the result.

(f) *Credits: 2p.* Let $T(y) = \max_{y_i}$. Calculate the posterior predictive p-value, $\Pr(T(y^{rep}) \geq T(y))$, by simulating from the posterior predictive density $p(T(y^{rep})|y)$ using the model and data in this problem 2. How well can the model replicate the length of the largest mollusc in this data?

3. Derivations and comparing posterior distributions (*Total credits: 17p*)

Problems (a), (b) and (c) should only be solved on **Paper**. Let $x|\theta$, where $x = (x_1, \ldots, x_n)$, be an independent sample from a Weibull distribution with the following density:

$$f(x_i|\theta) = 3\theta x_i^2 e^{-\theta x_i^3},\ x_i > 0, \theta > 0\ i = 1, \ldots, n.$$

Assume the following conjugate gamma prior distribution for θ: $p(\theta) = 32\theta^{3-1}e^{-4\theta}$.

(a) *Credits: 3p.* Derive the posterior distribution for θ.

(b) *Credits: 2p.* Compute the Bayes point estimate of θ for $x_1 = 0.8, x_2 = 1.1, x_3 = 0.8, x_4 = 0.9, x_5 = 1.0$, by assuming the zero-one loss function.

(c) *Credits: 5p.* Compute the marginal likelihood for the independent sample in 3(b) using the Laplace approximation. The Laplace approximation of the marginal likelihood is given by

$$\ln \hat{p}(x) = \ln p(x|\hat{\theta}) + \ln p(\hat{\theta}) + \frac{1}{2}\ln\left|J_{\hat{\theta},x}^{-1}\right| + \frac{1}{2}\ln(2\pi),$$

where $\hat{\theta}$ is the posterior mode and $J_{\hat{\theta},x} = -\frac{\partial^2 \ln p(\theta|x)}{\partial \theta^2}|_{\theta=\hat{\theta}}$ is the observed information at the posterior mode.

(d) *Credits: 4p.* Write a function in R that computes the log posterior distribution of θ for an independent sample $x|\theta$. Then, use this function to plot the posterior distribution of θ based on the data in 3(b).

(e) *Credits: 3p.* Use numerical optimization to obtain a normal approximation to the posterior distribution of θ based on the data in 3(b). Use the `lines` command in R to plot this approximate posterior in the same graph as the posterior obtained in 3(d). [Hints: use the argument `lower=0.1` in `optim`, and `method=c("L-BFGS-B")`]. Is the posterior approximation accurate?

Good luck!
Best, Bertil