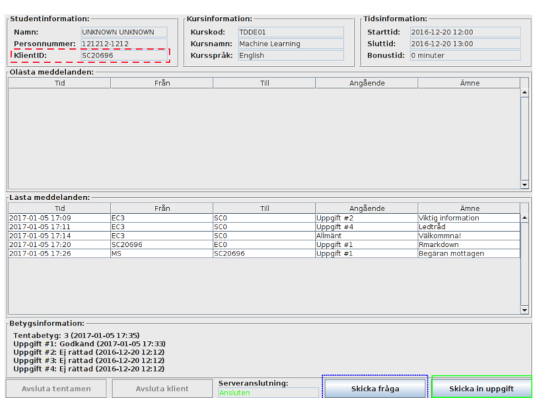
LINKÖPING UNIVERSITY
Dept. of Computer and Information Science
Division of Statistics and Machine Learning
Bertil Wegmann

2022-08-17
Bayesian Learning, 6 hp
732A73/732A91/TDDE07

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

Time:	14-18
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. Derivations and comparing posterior distributions (*Total credits: 20p*)

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

$$f(x_i|\theta) = \frac{\theta^3}{2x_i^4} e^{-\theta/x_i}, \ \theta > 0, \ x_i > 0, i = 1, \dots, n.$$

Assume the following gamma prior distribution for θ with parameters $\alpha = 4$ and $\beta = 2$: $p(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \theta^{\alpha-1} e^{-\beta\theta}$, where $\Gamma(\alpha) = (\alpha - 1)!$ for any positive integer α.

(a) *Credits: 4p.* Show that $p(\theta)$ is a conjugate prior distribution for θ by showing that the posterior distribution of θ is Gamma $\left(4 + 3n, 2 + \sum_{i=1}^{n} \frac{1}{x_i}\right)$.

(b) *Credits: 4p.* Show that the marginal likelihood for an independent sample $x|\theta$ from the model above is equal to

$$\frac{\Gamma(4+3n)2^4}{\Gamma(4)\left(2 + \sum_{i=1}^{n} \frac{1}{x_i}\right)^{4+3n} 2^n \prod_{i=1}^{n} x_i^4}$$

and calculate the exact value of the marginal likelihood for $x_1 = 0.7, x_2 = 1.1, x_3 = 0.9, x_4 = 1.5$.

(c) *Credits: 5p.* Compute the marginal likelihood for the independent sample in 1(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. Is the Laplace approximation of the marginal likelihood accurate?

(d) *Credits: 4p.* Write a function in R that computes the log posterior distribution of θ as the sum of the log likelihood and the log prior 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 1(b).

(e) *Credits: 3p.* Use numerical optimization to obtain a normal approximation to the posterior distribution of θ based on the data in 1(b). Use the `lines` command in R to plot this approximate posterior in the same graph as the posterior obtained in 1(d). [Hints: use the argument `lower=0.1` in `optim`, and `method=c("L-BFGS-B")`]. Is the posterior approximation accurate?

2. Logistic regression (*Total credits: 10p*)

The file `Bridge.RData`, which is loaded by the code in `ExamData.R`, contains data on 42 bridges. For each bridge i we have observed $y_i = 1$ *if bridge i needs to be repaired within the next five years* and 0 otherwise, $x_{1i} =$ *the number of years since the construction of bridge i* and $x_{2i} = 1$ *if bridge i was built using a certain technology* and 0 otherwise. The data matrix also contains a column *const* with ones to include an intercept in the model. Now, use `BayesLogitReg.R` in `ExamData.R` to sample from the joint posterior distribution of the model parameters in the following logistic regression model:

$$\Pr(y_i = 1 \mid x_1, x_2) = p_i = \frac{\exp(\beta_0 + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i})}{1 + \exp(\beta_0 + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i})}$$

Simulate 20000 draws from the joint posterior. Use the following prior distribution for $\beta = (\beta_0, \beta_1, \beta_2)$:

$$\beta \sim N(0, 10^2 I_3)$$

(a) *Credits: 1p.* Compute a 95% equal tail credible interval for β_1 and interpret it.

(b) *Credits: 1p.* Compute the joint posterior probability that both $\beta_1 > 0$ and $\beta_2 > 0$.

(c) *Credits: 2p.* Plot the posterior distribution of $\frac{1-p_i}{p_i}$ for a five-year-old bridge that was built with the certain technology. Does the plot seem reasonable? Should we in this problem question the reliability of results for such bridges? Motivate your answers.

(d) *Credits: 3p.* Consider a bridge i that was not built with the certain technology. Compute 95 % equal tail posterior probability intervals for p_i on a grid of values of x_1. 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.

(e) *Credits: 3p.* Consider a 40-year-old bridge i that was built using the certain technology. Plot the posterior distribution of p_i and compute the posterior probability that $p_i > 0.5$ for this bridge.

3. DECISION ON BUYING A FINANCIAL OPTION OR NOT (*Total credits: 10p*)

Problems (a) and (b) should only be solved on **Paper**. An investment bank has to make a decision every day whether or not to buy a certain financial option, in order to insure a stock portfolio against financial risks. The bank's utility depends on whether the portfolio value goes up or not and different utilities are given in the table below. Let $x \sim \text{Bern}(\theta)$ denote the portfolio outcome, where $x = 1$ if the portfolio value goes up and 0 otherwise.

	Portfolio value goes up	Portfolio value does not go up
Buy the option	60	-20
Don't buy the option	180	-240

(a) *Credits: 2p.* Assume that θ is known and equal to $\theta = 0.6$. Compute the Bayesian decision, whether the bank should buy the option or not.

(b) *Credits: 5p.* Now assume that θ is unknown, and that there are 50 independent observations $x_1, ..., x_{50} \overset{iid}{\sim} \text{Bern}(\theta)$ from previous days. Assume that the bank has a Beta$(6,4)$ prior for θ. Derive the predictive distribution for x_{51} on **Paper**. Assume that the portfolio value went up in 32 of the 50 previous days.
[Hints: It may be used that the posterior distribution for θ is Beta $(\alpha + s, \beta + f)$-distributed, with α and β being prior parameters and s and f being the number of observed successes and failures in the Bernoulli trials. The Gamma function has the property $\Gamma(y+1) = y\Gamma(y)$.]

(c) *Credits: 3p.* Compute the Bayesian decision for day 51 based on the information in (b). Solutions based on simulation and solutions based on the answer in (b) are both accepted.

GOOD LUCK!
BEST, BERTIL