

Bayesian Learning
Computer Lab 1

You are recommended to use R for solving the labs.

You work and submit your labs in pairs, but both of you should contribute equally and understand all parts of your solutions.

It is not allowed to share exact solutions with other student pairs.

The submitted lab reports will be verified through OURIGINAL and indications of plagiarism will be investigated by the Disciplinary Board.

Submit your solutions via LISAM, no later than April 17 at 23:59.

Please note the following about the format of the submitted lab report:

1. The lab report should include all solutions and plots to the stated problems with necessary comments.
 2. Submit the lab report with your code attached to the solution of each sub-problem (1a), (1b),...) in **one** PDF document.
 3. Submit a separate file containing all code.
-

1. *Daniel Bernoulli*

Let $y_1, \dots, y_n | \theta \sim \text{Bern}(\theta)$, and assume that you have obtained a sample with $s = 22$ successes in $n = 70$ trials. Assume a $\text{Beta}(\alpha_0, \beta_0)$ prior for θ and let $\alpha_0 = \beta_0 = 8$.

- (a) Draw 10000 random values (`nDraws = 10000`) from the posterior $\theta | y \sim \text{Beta}(\alpha_0 + s, \beta_0 + f)$, where $y = (y_1, \dots, y_n)$, and verify graphically that the posterior mean $E[\theta | y]$ and standard deviation $SD[\theta | y]$ converges to the true values as the number of random draws grows large. [Hint: `rbeta()` to draw random values and make graphs of the sample means and standard deviations of θ as a function of the accumulating number of drawn values].
- (b) Draw 10000 random values from the posterior to compute the posterior probability $\Pr(\theta > 0.3 | y)$ and compare with the exact value from the Beta posterior. [Hint: `pbeta()`].
- (c) Draw 10000 random values from the posterior of the odds $\phi = \frac{\theta}{1-\theta}$ by using the previous random draws from the Beta posterior for θ and plot the posterior distribution of ϕ . [Hint: `hist()` and `density()` can be utilized].

2. *Log-normal distribution and the Gini coefficient.*

Assume that you have asked 8 randomly selected persons about their monthly income (in thousands Swedish Krona) and obtained the following eight observations: 33, 24, 48, 32, 55, 74, 23, and 17. A common model for non-negative continuous

variables is the log-normal distribution. The log-normal distribution $\log \mathcal{N}(\mu, \sigma^2)$ has density function

$$p(y|\mu, \sigma^2) = \frac{1}{y \cdot \sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2\sigma^2} (\log y - \mu)^2 \right],$$

where $y > 0$, $-\infty < \mu < \infty$ and $\sigma^2 > 0$. The log-normal distribution is related to the normal distribution as follows: if $y \sim \log \mathcal{N}(\mu, \sigma^2)$ then $\log y \sim \mathcal{N}(\mu, \sigma^2)$. Let $y_1, \dots, y_n | \mu, \sigma^2 \stackrel{iid}{\sim} \log \mathcal{N}(\mu, \sigma^2)$, where $\mu = 3.6$ is assumed to be known but σ^2 is unknown with non-informative prior $p(\sigma^2) \propto 1/\sigma^2$. The posterior for σ^2 is the $Inv - \chi^2(n, \tau^2)$ distribution, where

$$\tau^2 = \frac{\sum_{i=1}^n (\log y_i - \mu)^2}{n}.$$

- (a) Draw 10000 random values from the posterior of σ^2 by assuming $\mu = 3.6$ and plot the posterior distribution.
 - (b) The most common measure of income inequality is the Gini coefficient, G , where $0 \leq G \leq 1$. $G = 0$ means a completely equal income distribution, whereas $G = 1$ means complete income inequality (see e.g. Wikipedia for more information about the Gini coefficient). It can be shown that $G = 2\Phi(\sigma/\sqrt{2}) - 1$ when incomes follow a $\log \mathcal{N}(\mu, \sigma^2)$ distribution. $\Phi(z)$ is the cumulative distribution function (CDF) for the standard normal distribution with mean zero and unit variance. Use the posterior draws in a) to compute the posterior distribution of the Gini coefficient G for the current data set.
 - (c) Use the posterior draws from b) to compute a 95% equal tail credible interval for G . A 95% equal tail credible interval (a, b) cuts off 2.5% percent of the posterior probability mass to the left of a , and 2.5% to the right of b .
 - (d) Use the posterior draws from b) to compute a 95% Highest Posterior Density Interval (HPDI) for G . Compare the two intervals in (c) and (d). [Hint: do a kernel density estimate of the posterior of G using the `density` function in R with default settings, and use that kernel density estimate to compute the HPDI. Note that you need to order/sort the estimated density values to obtain the HPDI.].
3. *Bayesian inference for the concentration parameter in the von Mises distribution.*
This exercise is concerned with directional data. The point is to show you that the posterior distribution for somewhat weird models can be obtained by plotting it over a grid of values. The data points are observed wind directions at a given location on ten different days. The data are recorded in degrees:

$$(20, 314, 285, 40, 308, 314, 299, 296, 303, 326),$$

where North is located at zero degrees (see Figure 1 on the next page, where the angles are measured clockwise). To fit with Wikipedia's description of probability distributions for circular data we convert the data into radians $-\pi \leq y \leq \pi$. The 10 observations in radians are

$$(-2.79, 2.33, 1.83, -2.44, 2.23, 2.33, 2.07, 2.02, 2.14, 2.54).$$

Assume that these data points conditional on (μ, κ) are independent observations from the following von Mises distribution:

$$p(y|\mu, \kappa) = \frac{\exp[\kappa \cdot \cos(y - \mu)]}{2\pi I_0(\kappa)}, \quad -\pi \leq y \leq \pi,$$

where $I_0(\kappa)$ is the modified Bessel function of the first kind of order zero [see `?besselI` in R]. The parameter μ ($-\pi \leq \mu \leq \pi$) is the mean direction and $\kappa > 0$ is called the concentration parameter. Large κ gives a small variance around μ , and vice versa. Assume that μ is known to be 2.4. Let $\kappa \sim \text{Exponential}(\lambda = 0.5)$ a priori, where λ is the rate parameter of the exponential distribution (so that the mean is $1/\lambda$).

- (a) Derive the expression for what the posterior $p(\kappa|y, \mu)$ is proportional to. Hence, derive the function $f(\kappa)$ such that $p(\kappa|y, \mu) \propto f(\kappa)$. Then, plot the posterior distribution of κ for the wind direction data over a fine grid of κ values. [Hint: you need to normalize the posterior distribution of κ so that it integrates to one.]
- (b) Find the (approximate) posterior mode of κ from the information in a).

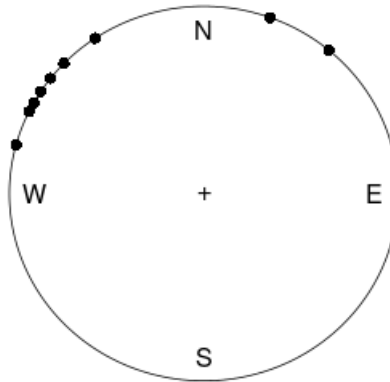


Figure 1: The wind direction data. Angles are measured clock-wise starting from North.

GOOD LUCK!