S2: Statistical Methods for Data Science

Example Sheet 1

The value of a quantity X is measured N times and the following measurement values obtained:

$$\{x_i\} = \{x_1, x_2, \ldots, x_N\} \,.$$

The errors on each measurement are independent and Gaussian distributed with a known standard deviation of σ.

1. Write down an expression for the likelihood $\mathcal{L}(\{x_i\}|X)$.

2. Use a uniform prior distribution $\pi(X)$ and ensure that the prior is properly normalised. Write down Bayes' theorem for this problem and find an expression for the posterior $P(X|\{x_i\})$.

3. Write down an integral expression for the Bayesian evidence, Z. You do not need to evaluate this integral.

4. Write down expressions for (i) the most likely value of the parameter X (i.e. the maximum *a posteriori* value), (ii) the mean of the posterior $P(X|\{x_i\})$, and (iii) the standard deviation of the posterior $P(X|\{x_i\})$. To what extent do these quantities depend on the prior bounds? Evaluate all three quantities that the prior bounds tend to $\pm\infty$.

5. Now imagine repeating the above analysis using a different prior, $\pi(X)$. You are free to choose the prior. Can you find a parameterisation for the prior such that the prior and the posterior are in the same parametric family of distributions. In other words, find the conjugate

prior for X for this likelihood function. Write the Bayesian inference in the notation used in Box 1.6 (equations i, ii & iii) of the lecture notes. State explicitly the rules for updating the parameters of the prior/posterior distribution.

6. Find the Bayesian evidence for your conjugate prior in part 5. *[Hint: you do not need to do any integration for this. But the algebra can get a bit messy. It is easiest to give you answer in terms of N, σ, μ, μ', Σ, Σ' and x_i.]*

The inverse gamma distribution has a probability density function that is defined as

$$f(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} \exp\left(\frac{-\beta}{x}\right),$$

for $x > 0$ and with shape and scale parameters $\alpha > 0$ and $\beta > 0$ respectively.

7. Now suppose that the value of parameter X is known but the variance in the data σ^2 is unknown. Working from the same data $\{x_i\}$, our task is now to infer the value σ^2. Write down the likelihood and a new version of Bayes' theorem for this problem. Show that $\pi(\sigma^2) = f(\sigma^2|\alpha,\beta)$ is a conjugate prior for this problem and find the rule for updating the parameters of the inverse gamma prior/posterior distribution.

8. If both X and σ^2 are unknown and have to be inferred simultaneously, can you suggest a suitable two-dimensional conjugate prior $\pi(X,\sigma^2)$? Compare your suggestion with the definition of the *normal-inverse-gamma distribution* (see, e.g. www.wikipedia.org/wiki/Normal-inverse-gamma_distribution). Is this prior separable?