

UNIVERSITETET I OSLO

Matematisk Institutt

EXAM IN: **STK 4021/9021 – Applied Bayesian Analysis**
WITH: **Nils Lid Hjort**
AUXILIA: **Calculator, plus one single sheet of paper
with the candidate's own personal notes**
TIME FOR EXAM: **Tuesday 4/Dec/2025, 9:00 – 13:00**

This exam set contains four exercises and comprises four pages, including a brief appendix on the gamma distribution.

Exercise 1: Theory Zero versus Theory One

IN THEORY THERE IS NO DIFFERENCE between theory and practice; in practice there is. Here we work with methods able to spot even a tiny difference, with enough data.



The probabilities p_n (red, slanted) and q_n (black, full), for Theories Zero ($a = 1.00$) and One ($a = 1.25$), as a function of n , for data simulated from Theory One.

- (a) Assume $Y | \theta$ is an $\text{Expo}(\theta)$, with density $\theta \exp(-\theta y)$ on the positive halfline, and mean $1/\theta$. Suppose further that the background parameter θ has been drawn from the gamma (a, b) distribution (see the Appendix for the formula). Find the posterior distribution of θ given $Y = y$.

- (b) Show that the marginal density is

$$f(y, a, b) = \frac{ab^a}{(b+y)^{a+1}} \quad \text{for } y > 0.$$

What is its mean?

- (c) Independent identically distributed data $Y_1, Y_2, \dots$ come in your door, one by one. In general terms, these are either of Type Zero, with density $f_0(y)$, or of Type One, with density $f_1(y)$, with prior probabilities $\frac{1}{2}, \frac{1}{2}$ for these two possibilities. Show that after n datapoints, we have

$$p_n = \Pr(\text{Zero} | y_1, \dots, y_n) = \frac{1}{1 + W_n}, \quad q_n = \Pr(\text{One} | y_1, \dots, y_n) = \frac{W_n}{1 + W_n},$$

in which $W_n = \prod_{i=1}^n f_1(y_i)/f_0(y_i)$.

- (d) Now apply this machinery to the case where the Y_i are as in (a) and (b), each Y_i the result of the two-step random mechanism with θ_i a gamma $(a, 1)$ and then $Y_i | \theta_i \sim \text{Expo}(\theta_i)$ (so $b = 1$ in the formulation above). Theories Zero and One agree that data are of this type, but Zero holds that $a = 1.00$ and One believes that $a = 1.25$. First, for a single data point y_1 , for which interval will Theory One look stronger than Theory Zero?
- (e) Then write out what the above general formulae imply for p_n and q_n , in this setup. Comment on the figure above. Show also that if data really come from One, then $q_n \rightarrow 1$, with probability one.

Exercise 2: skewing the uniform

I HAVE A PERFECTLY AVERAGE SKEWED PERCEPTION of myself. For unit interval data being perhaps skewed compared to the uniform, consider the density

$$f(y, \gamma) = \gamma y^{\gamma-1} \quad \text{for } y \in (0, 1),$$

with γ a positive unknown parameter; close to uniformity means γ not far from 1.

- (a) Find a formula for its median, $\mu(\gamma)$. Find also the Fisher information $J(\gamma)$ and the Jeffreys prior, and comment briefly on this.
- (b) Suppose $Y_1, \dots, Y_n$ are independent observations from the model above. Write down an expression for the likelihood function, and show that the maximum likelihood estimator becomes

$$\hat{\gamma}_{\text{ml}} = \frac{1}{W_n}, \quad \text{with } W_n = \frac{1}{n} \sum_{i=1}^n \log(1/Y_i) = -\frac{1}{n} \sum_{i=1}^n \log Y_i.$$

- (c) With a gamma (a, b) prior for γ , for some prior elicited values of (a, b) , show that the posterior is another gamma, with updated parameters. Show in particular that the posterior mean Bayes estimator is

$$\hat{\gamma}_B = \frac{a + n}{b + nW_n}.$$

- (d) Explain how you can find a 90 percent credibility interval for the median $\mu(\gamma)$, for the given dataset and this given prior.
- (e) Comment on how ‘statistically close’ the maximum likelihood and the Bayes estimator are.
- (f) In cases where it is not easy to elicit gamma prior parameters (a, b) , what could perhaps be done?

Exercise 3: copulas and copulae

IT’S COMPLICATED. Consider the coupling model for (U, V) on the unit square with joint density

$$f(u, v, \theta) = 1 + \theta(u - \frac{1}{2})(v - \frac{1}{2}),$$

with θ the copula parameter.

- (a) What is the range of θ that makes this a density on the $[0, 1] \times [0, 1]$? Show that the marginals are uniform.
- (b) Suppose you are given 33 pairs (u_i, v_i) drawn from this density, and use the prior $\pi(\theta)$ which, for this occasion, is flat on $[0, 2]$. Give a formula for the posterior $\pi(\theta | \text{data})$. In this relatively simple setup there is no real need for ‘sophisticated computing’, but explain, with a few but not many details, how you can set up an MCMC for sampling from this posterior distribution.

Exercise 4: estimating two normal means

THE END JUSTIFIES THE MEANS, so let’s estimate two of them. The setup is that $X \sim N(a, 1)$, $Y \sim N(b, 1)$, and that these are independent. We write their densities in the usual form $\phi(x - a)$ and $\phi(y - b)$, with $\phi(\cdot)$ the standard normal density. The loss function to be used is

$$L(a, b, \hat{a}, \hat{b}) = (\hat{a} - a)^2 + (\hat{b} - b)^2.$$

- (a) With the simple natural estimator (X, Y) itself, find the risk function.
- (b) Show that the Jeffreys prior for (a, b) is the flat one, say $\pi_1(a, b) = 1$, constant over the full plane. With this prior, show that $a | (x, y) \sim N(x, 1)$ and $b | (x, y) \sim N(y, 1)$. Find the Bayes estimators $(\hat{a}, \hat{b})$.
- (c) A rather different prior is betting that the two are equal, and then with the common $a = b$ being flat on the line, i.e. $\pi_2(a, b) = 1$ on the diagonal $a = b$ and zero elsewhere. Show that $a | (x, y) \sim N(\frac{1}{2}(x + y), \frac{1}{2})$. For the estimator $(\tilde{a}, \tilde{b}) = (\frac{1}{2}(x + y), \frac{1}{2}(x + y))$, find its risk function, and comment.

- (d) Consider then the mixture prior $\pi^* = \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2$. Explain in words what this prior view says. Below, let

$$\bar{f}_1 = \int \int \phi(x-a)\phi(y-b) da db, \quad \bar{f}_2 = \int \phi(x-a)\phi(y-a) da$$

be the marginal distributions implied by the two priors, computed at the observed value (x, y) . Show that the posterior can be represented as

$$\pi^*(a, b | x, y) = p_1^* \pi_1(a, b | x, y) + p_2^* \pi_2(a, b | x, y),$$

in terms of the two posteriors already found, and with

$$p_1^* = \frac{\bar{f}_1}{\bar{f}_1 + \bar{f}_2}, \quad p_2^* = \frac{\bar{f}_2}{\bar{f}_1 + \bar{f}_2}.$$

- (e) Show from this that the Bayes estimators become

$$a^* = p_1^* X + p_2^* \frac{1}{2}(X + Y), \quad b^* = p_1^* Y + p_2^* \frac{1}{2}(X + Y),$$

and comment briefly on their form.

- (f) Find expressions for $\bar{f}_1$ and $\bar{f}_2$, and hence p_1^* and p_2^* , to complete the description of this mixture-Bayes way of estimating (a, b) .
- (g) One may show (but not now, during firetimerseksamen) that the risk function

$$R(a, b, a^*, b^*)$$

depends only on $d = b - a$, the mean difference. Draw a plot with two curves, as functions of d : the first is for the standard estimator (X, Y) , and the second one is meant to be your informed guess, or rough sketch, for the (a^*, b^*) method. One would need a computer to actually do this with precision (e.g. via simulations at each grid point d), but the idea is that you should use your intuition and insights to provide such a rough sketch.

Appendix: just some basis facts for the Gamma distribution

We say that X has the Gamma distribution with parameters (a, b) if its density is

$$f(x, a, b) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx) \quad \text{for } x > 0.$$

Its mean and variance are a/b and a/b^2 . Also, the inverse mean is $b/(a-1)$ for $a > 1$.