Exam Statistical Reasoning

Date: Friday, November 9, 2018
Time: 09.00-12.00
Place: MartiniPlaza, Groningen
Progress code: WISR-11

Rules to follow:

- This is a closed book exam. Consultation of books and notes is **not** permitted.
- Do not forget to write your name and student number onto each paper sheet and onto the envelope.
- There are 4 exercises and the number of points per exercise are indicated within boxes. You can reach 90 points.
- If you have to compute something, then include the relevant equations, so that it is clear how you obtained the result.
- **We wish you success with the completion of the exam!**

START OF EXAM

1. **Negative binomial distribution with Gamma prior.** [20]
   The random variables $Y_1, \dots, Y_n$ are negative binomial distributed with the two parameters $r \in \mathbb{N}$ and $\theta \in (0,1)$. Assume that r is known, while θ is unknown. We impose a Beta prior with the hyperparameters $a > 0$ and $b > 0$ on θ.

   (a) [10] Compute the posterior distribution of θ.

   (b) [5] For the special case: $r = 1$, interpret the hyperparameters a and b in terms of pseudo counts.

   (c) [5] For the special case: $r = 1$, $a = 1$, $b = 1$, compute the marginal likelihood.

   <u>Recall</u>:
   The density of the negative binomial distribution with $r \in \mathbb{N}$ and $\theta \in [0,1]$ is

   $$p(x|\theta, r) = \binom{r+x-1}{x} \cdot \theta^r \cdot (1-\theta)^x \qquad (x \in \mathbb{N}_0)$$

   The density of a Beta distribution with parameters $a > 0$ and $b > 0$ is:

   $$p(x|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot x^{a-1} \cdot (1-x)^{b-1} \qquad (0 < x < 1)$$

2. **Gamma distribution with exponential prior.** $\boxed{\textbf{30}}$

Assume that the random variables $Y_1, \ldots, Y_n$ have a Gamma distribution:

$$Y_1, \ldots, Y_n | (a, b) \sim \text{GAM}(a, b)$$

and that $Y_1, \ldots, Y_n$ are independent given the two parameters $a > 0$ and $b > 0$. Assume that both parameters a and b are unknown. The prior of b is an exponential distribution with parameter δ. The prior of a is an exponential distribution with parameter λ. Assume that the data $Y_1 = y_1, \ldots, Y_n = y_n$ have been observed.

(a) $\boxed{10}$ Compute the full conditional distribution of b.

Now assume that $a = 1$ is known, so that the model has only one single unknown parameter b with an exponential prior. We set the hyperparameter δ equal to 2, and we have observed $n = 5$ realisations: $y_1 = 3$, $y_2 = 2$, $y_3 = 0.5$, $y_4 = 2$, and $y_5 = 0.5$.

(b) $\boxed{10}$ Compute the marginal likelihood.

(c) $\boxed{10}$ Compute the predictive probability for $\tilde{Y} = 2$.

<u>HINTS</u>:

The density of a Gamma distribution with parameters $a > 0$ and $b > 0$ is:

$$p(x|a, b) = \frac{b^a}{\Gamma(a)} \cdot x^{a-1} \cdot e^{-b \cdot x} \qquad (x > 0)$$

The density of the exponential distribution with parameter $\lambda > 0$ is

$$p(x|\lambda) = \lambda \cdot e^{-\lambda \cdot x} \qquad (x > 0)$$

3. **Markov Chain Monte Carlo (MCMC) pseudo code.** $\boxed{\textbf{10}}$

Consider a model with n variables $Y_1, \ldots, Y_n$, whose distribution depends on a parameter $\theta \in \mathbb{R}$. The variables $Y_1, \ldots, Y_n$ are i.i.d. conditional on θ and have the density $p(y|\theta)$. The parameter θ is unknown and has a non-conjugate prior distribution with density $p(\theta)$. Let $y_1, \ldots, y_n$ be the observed data.

Describe in terms of pseudo code how a (Metropolis-Hastings) Markov Chain Monte Carlo simulation can be used to generate a sample from the posterior distribution. <u>Hint</u>: You might want to complete the following:

<u>Initialisation</u>: Set $\theta^{(0)} = \theta$, where $\theta \in \mathbb{R}$.

<u>Iterations</u> For $t = 1, \ldots, T$

- Sample ...
- Propose ...
- Compute the acceptance probability $A(.,.) = \ldots$
- Sample ...
- IF ... ELSE ...

<u>Output</u>: $\theta^{(0)}, \ldots, \theta^{(T)}$

4. **Gaussian and Gamma distributions.** $\boxed{30}$

(a) $\boxed{5}$ Assume that the n-dimensional vector $\mathbf{x}$ has a Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Show for the density of $\mathbf{x}$:

$$p(\mathbf{x}) \propto \exp\{-\frac{1}{2} \cdot \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \cdot \mathbf{x} + \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}\}$$

(b) $\boxed{10}$ Show that the n-dimensional Gaussian with $\boldsymbol{\mu} = \mu \cdot \mathbf{1}$ and $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$, where $\mathbf{1}$ is a vector of ones and $\mathbf{I}$ is the identity matrix, implies that the density at $\mathbf{x} = (x_1, \ldots, x_n)^T$ is equal to the joint density of an i.i.d. sample $x_1, \ldots, x_n$ from a one-dimensional $N(\mu, \sigma^2)$ distribution.

(c) $\boxed{5}$ Assume that the density of a random variable X with sample space $\mathbb{R}$ fulfills:

$$p(x) \propto e^{-2x^2 - 4x + 7}$$

What is then the distribution of X?

(d) $\boxed{5}$ Assume that the density of a random variable X with sample space $\mathbb{R}^+$ fulfills:

$$p(x) \propto e^{-4x + 7}$$

What is then the distribution of X?

(e) $\boxed{5}$ Assume that the density of a random variable λ with sample space $\mathbb{R}^+$ is proportional to the density of the n-dimensional Gaussian with mean vector $\boldsymbol{\mu} = \mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma} = \lambda^{-1} \mathbf{W}$.
What is then the distribution of λ?

RECALL:
The n-dimensional Gaussian with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ has the density:

$$p(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-n/2} \cdot \det(\boldsymbol{\Sigma})^{-1/2} \cdot \exp\{-\frac{1}{2} \cdot (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\} \qquad (\mathbf{x} \in \mathbb{R}^n)$$

The 1-dimensional Gaussian with mean μ and variance σ^2 has the density:

$$p(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma} \cdot \exp\{-\frac{1}{2} \cdot \frac{(x - \mu)^2}{\sigma^2}\} \qquad (x \in \mathbb{R})$$

The density of a Gamma distribution with parameters $a > 0$ and $b > 0$ is:

$$p(x|a, b) = \frac{b^a}{\Gamma(a)} \cdot x^{a-1} \cdot e^{-b \cdot x} \qquad (x > 0)$$

For an n-by-n matrix A and $\alpha \in \mathbb{R}$ it holds:

$$\det(\alpha \cdot A) = \alpha^n \cdot \det(A)$$