Exam Statistical Reasoning

Date: Friday, November 6, 2015
Time: 09.00-12.00
Place: A.Jacobshal 01
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 fill in your name and student number.
- The number of points per question are indicated within a box. Ten points are free.
- We wish you success with the completion of the exam!

START OF EXAM

1. **Two Bayes Theorem exercises.** $\boxed{10}$

   (a) $\boxed{5}$ There are two urns containing colored balls. The first urn contains 50 red balls and 50 blue balls. The second urn contains 30 red balls and 70 blue balls. One of the two urns is randomly chosen (both urns have probability $p = 0.5$ of being chosen) and then a ball is drawn at random from one of the two urns. If a red ball is drawn, what is the probability that it comes from the first urn?

   (b) $\boxed{5}$ Alice has two coins in her pocket, a fair coin (head on one side and tail on the other side) and a two-headed coin. She picks one at random from her pocket (both coins have probability $p = 0.5$ of being chosen), tosses it and obtains head. What is the probability that she flipped the fair coin?

2. **Density of the Gaussian distribution.** $\boxed{15}$

   (a) $\boxed{10}$ The density (pdf) of a Gaussian distribution with parameters μ and σ^2:

   $$p(x) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma} \cdot \exp\{-0.5 \cdot \frac{(x-\mu)^2}{\sigma^2}\}$$

   can be written as:

   $$p(x) = c \cdot \exp\{-0.5 \cdot a \cdot x^2 + b \cdot x\}$$

   where a, b and c do **not** depend on x. Give the constants a, b and c.

   (b) $\boxed{5}$ In a Bayesian analysis it has been shown that the posterior density (pdf) of a parameter θ is proportional to:

   $$p(\theta|y_1, \ldots, y_n) \propto \exp\{-0.25 \cdot \theta^2 + 0.5 \cdot \theta\}$$

   What is the posterior distribution of θ?

to be continued below

3. **Exponential-Gamma model.** $\boxed{20}$
Consider a set of n continuous random variables $Y_1, \ldots, Y_n$ which are i.i.d. exponentially distributed with rate parameter $\lambda > 0$, symbolically:

$$Y_1, ..., Y_n|\lambda \sim Exp(\lambda)$$

For exponential distributions the set of potential outcomes is $\mathbb{R}_0^+$ and each individual variable Y_i has the following probability density function (pdf). For $y_i \geq 0$:

$$p(y_i|\lambda) = \lambda \cdot e^{-\lambda \cdot y_i}$$

(a) $\boxed{2}$ Show that the exponential distribution is effectively just a special case of the Gamma distribution. Recall that the density (pdf) of a Gamma distribution with parameters α and β is the following one. For all $x > 0$:

$$p(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot x^{\alpha-1} \cdot e^{-\beta \cdot x}$$

(b) $\boxed{3}$ Compute the joint probability density function $p(y_1, \ldots, y_n|\lambda)$ of the n exponentially distributed variables $Y_1, ..., Y_n$.

(c) $\boxed{10}$ Assume that the realisations $Y_1 = y_1, \ldots, Y_n = y_n$ have been observed. Impose a Gamma-Prior with parameters α and β on the unknown rate parameter λ and show that the posterior distribution of λ is then a Gamma distribution with parameteters $\tilde{\alpha} := \alpha + n$ and $\tilde{\beta} := \beta + \sum_{i=1}^n y_i$.
Note that the density (pdf) of a Gamma distribution is given in part (b).

(d) $\boxed{5}$ What is the interpretation of the hyperparameters α and β of the Gamma prior in terms of pseudo-counts?

4. **Predictive distribution of the Exponential-Gamma model.** $\boxed{25}$
Re-consider the Exponential-Gamma model from exercise 2.

(a) $\boxed{15}$ Compute the predictive distribution of a new random variable $\tilde{Y}$, where $Y_1, \ldots, Y_n$ and $\tilde{Y}$ are i.i.d conditional on λ.

<u>HINT</u>: A continuous random variable Z is Lomax-distributed with parameters $a > 0$ and $b > 0$ if its density (PDF) is given by:

$$p(z|a, b) = \frac{a}{b}(1 + \frac{z}{b})^{-(a+1)}$$

for all $z \in \mathbb{R}_0^+$.

(b) $\boxed{10}$ Assume that there is no way to generate random samples from the Lomax distribution, while realisations of standard distributions, such as Gamma- and Exponential-distributions, can be generated straightforwardly.
Describe in terms of pseudo code how the Monte-Carlo method can be employed to approximate the expectation $E[\tilde{Y}|Y_1 = y_1, \ldots, Y_N = y_n]$ of the predictive Lomax distribution from part (a).

to be continued below

5. **Discrete Markov chains** $\boxed{15}$
Consider a simple discrete random variable X with sample space $\Theta = \{1, 2, 3\}$ and discrete density (pdf): $p(1) = 0.5$, $p(2) = 0.4$, and $p(3) = 0.1$. The goal is to define a Markov chain whose stationary distribution is the distribution of X. To this end, a Metropolis-Hastings MCMC sampling scheme can be employed.
Assume that the proposal probabilities are given by: $Q(1, 1) = 0$, $Q(1, 2) = 1$, $Q(1, 3) = 0$, $Q(2, 1) = 0.1$, $Q(2, 2) = 0$, $Q(2, 3) = 0.9$, $Q(3, 1) = 0$, $Q(3, 2) = 1$, and $Q(3, 3) = 0$, where $Q(i, j)$ is the probability to propose to move from state $i \in \Theta$ to state $j \in \Theta$.

   (a) $\boxed{5}$ Compute the Metropolis-Hastings acceptance probabilities $A(1, 2)$, $A(2, 1)$, $A(2, 3)$, and $A(3, 2)$, where $A(i, j)$ is the probability to accept a proposed move from state $i \in \Theta$ to state $j \in \Theta$.

   (b) $\boxed{10}$ Give the 3-by3 transition matrix of the resulting Markov chain.

6. **Graphical Models and full conditional distributions.** $\boxed{15}$
Consider the following hierarchical Bayesian model.
The sampling model is given by n variables $Y_1, \ldots, Y_n$ which are Bernoulli distributed, $Y_1, \ldots, Y_n|\theta \sim \text{Ber}(\theta)$, where $\theta \in [0, 1]$ is the unknown probability parameter and $Y_1, \ldots, Y_n$ are i.i.d. conditional on θ. A Beta prior with parameters $a > 0$ and $b > 0$ is imposed on θ, whose hyperparameters a and b are both unknown. Therefore two independent Gamma hyperpriors (with hyper-hyperparameters α_a and β_a and hyperparameters α_b and β_b) are imposed on a and b. Symbolically:

$$\begin{aligned} Y_1, \ldots, Y_n|\theta &\sim \text{Ber}(\theta) \\ \theta|a, b &\sim \text{Beta}(a, b) \\ a &\sim \text{Gamma}(\alpha_a, \beta_a) \\ b &\sim \text{Gamma}(\alpha_b, \beta_b) \end{aligned}$$

where $\alpha_a > 0$, $\beta_a > 0$, $\alpha_b > 0$, and $\beta_b > 0$, are fixed and known.

   (a) $\boxed{5}$ Give a graphical model representation of this hierarchical model.

   (b) $\boxed{5}$ Are the following statements right (R) or wrong (W):
   (i) The random variables $Y_1, \ldots, Y_n$ are (stochastically) independent.
   (ii) Given θ the random variables $Y_1, \ldots, Y_n$ are (stochastically) independent.
   (iii) The random variables a and b are (stochastically) independent.
   (iv) Given θ the random variables a and b are (stochastically) independent.
   (v) The random variables $Y_, \ldots, Y_n$ are (stochastically) independent of a.

   (c) $\boxed{5}$ Derive that the densities (pdfs) of the full conditional distributions fulfil:

$$\begin{aligned} p(\theta|a, b, y_1, \ldots, y_n) &\propto p(y_1, \ldots, y_n|\theta) \cdot p(\theta|a, b) \\ p(b|a, \theta, y_1, \ldots, y_n) &\propto p(\theta|a, b) \cdot p(b) \\ p(a|b, \theta, y_1, \ldots, y_n) &\propto p(\theta|a, b) \cdot p(a) \end{aligned}$$

   HINT: For *each* equality or proportionality that your derivations include give a brief explanation (e.g. single keywords). For example: '*by definition*' or '*as a function of x*', '*given y, x is independent of z*', etc.

END OF EXAM