Bayesian Statistics, Final Exam, MIMUW UW, January 2019

You can use notes and/or textbooks. Please do not communicate; do not use mobile phones or computers; do not cheat.

---

1. Let $X_1, \ldots, X_n$ be an i.i.d. sample from the normal distribution $\mathrm{N}(\theta, \sigma^2)$ with known σ. Consider the standard conjugate prior $\theta \sim \mathrm{N}(\mu, v^2)$.

   (a) Compute the shortest credible interval for θ at level $1 - \alpha = 0.95$, i.e. $[a, b]$ such that $\mathbb{P}(a \le \theta \le b | x_1, \ldots, x_n) = 1 - \alpha$.

   (b) What is the mninimum sample size n such that the length of this interval is bounded by a given d, i.e. $b - a \le d$?

   (c) Give the answer to the previous question in the case when $v = \infty$ (pass to the limit as $v \to \infty$).

2. As in the previous problem, $X_1, \ldots, X_n$ is an i.i.d. sample from $\mathrm{N}(\theta, \sigma^2)$ with known σ and $\theta \sim \mathrm{N}(\mu, v^2)$.

   (a) Give the (marginal) distribution of $\bar{X}$ in this model.

   (b) Give the (marginal) distribution of $X_1 - X_2$.

   (c) Compute $\mathrm{Cov}(\theta, \bar{X} - \theta)$.

   (d) Compute $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i^2$.

3. We observe a (1-dim)random variable X, which belongs to one of two classes: $C = 1$ or $C = 2$. Class-conditional probability distributions are normal,

   $(X|C = 1) \sim \mathrm{N}(-\mu, 1)$ and $(X|C = 2) \sim \mathrm{N}(\mu, 1)$.

   Prior probabilities of both the classes are equal, $\mathbb{P}(C = 1) = \mathbb{P}(C = 2) = 1/2$. We are to compute a “discriminant function” $\delta : \mathbb{R} \to \mathbb{R}$ which minimizes the Bayes risk $r(\delta) = \mathbb{E}\ell(C, \delta(X))$ for the following exponential loss function:

$$\ell(c, \delta(x)) = \begin{cases} \mathrm{e}^{-\delta(x)} & \text{for } c = 2; \\ \mathrm{e}^{\delta(x)} & \text{for } c = 1. \end{cases}$$

   (a) Compute the optimal function δ^*.

   (b) Compute the posterior risk $r_x(\delta^*)$ for this function.

   (c) Compute the Bayes risk $r(\delta^*)$ for this function in terms of μ.

   Hint: Express $\delta^*(x)$ in terms of the posterior probabilities $\mathbb{P}(C = 2|x)$ and $\mathbb{P}(C = 1|x)$.

4. Let $X|\theta \sim \text{Poiss}(\theta)$. We observe $X = x$ and consider the problem of testing

   $H_0 : \theta = \lambda$, where λ is a given positive number;

   against

   $H_1 : \theta \sim \text{Ex}(1/\lambda)$ (the exponential distribution with expectation λ).

   (a) Compute the Bayes Factor $B_{10}(x) = \mathbb{P}(H_1|x)/\mathbb{P}(H_0|x)$.
   (b) For what value of x the Bayes Factor $B_{10}(x)$ is minimal?
   (c) Consider the symmetric 0-1 loss and equal prior probabilities of the two hypotheses, $\mathbb{P}(H_0) = \mathbb{P}(H_1) = 1/2$. Let δ^* be the optimal test, i.e. the Bayes decision rule $\delta^* : \{0, 1, 2, \ldots\} \to \{0, 1\}$. What is the decision $\delta^*(x)$ in the special case $\lambda = 1$ and $x = 2$?

   Hint: To compute the Bayes Factor you can use the known formula for the marginal in the Poisson/Gamma model or directly integrate out θ. To find a minimum, compute $B_{10}(x)/B_{10}(x-1)$.

5. Let $(X_0, X_1, \ldots, X_k) = X_{0:k}$ be a Markov chain on the state space $\{1, 2\}$ with the transition matrix

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{pmatrix}.$$

   We assume that the chain is strictly stationary (the initial distribution $\mathbb{P}(X_0 = 1) = \mathbb{P}(X_0 = 2) = 1/2$ is stationary). We consider a Hidden Markov Model (HMM): variables $X_0, X_1, \ldots, X_k$ are not directly observed. We observe random variables $Y_0, Y_1, \ldots, Y_k$ with values in $\{1, 2\}$ such that Y_i depends only on X_i,

$$L(x, y) = \mathbb{P}(Y_i = y|X_i = x) = \begin{cases} 1-\varepsilon & \text{for } y = x; \\ \varepsilon & \text{for } y \neq x, \end{cases} \qquad x, y \in \{1, 2\}.$$

   We are interested in the posterior distribution $\pi_{\text{post}}(x_{0:k}) = \mathbb{P}(X_{0:k} = x_{0:k}|y_{0:k})$ on the space $\{1, 2\}^{k+1}$. To construct a Gibbs Sampler, we need full conditional distributions $\pi_{\text{post}}(\cdot)$, that is $\pi_{\text{post}}(x_i|x_{-i})$ where $x_{-i} = (x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k)$.

   (a) Compute $\mathbb{P}(X_i = 2|X_{i-1} = 2, X_{i+1} = 1, Y_i = 1)$ (for $0 < i < k$).
   (b) Consider other configurations (there are not so many essentially different cases).

6. Let $X_1, \ldots, X_n$ be an i.i.d. sample from the density

$$p_\theta(x) = \begin{cases} \theta x^{\theta-1} & \text{for } 0 < x < 1; \\ 0 & \text{otherwise.} \end{cases}$$

   The prior distribution of the parameter θ is $\text{Gamma}(\alpha, \lambda)$.

   (a) Find a 1-dim sufficient statistic in this model.
   (b) Compute the posterior distribution of $\theta|X_1, \ldots, X_n$.
   (c) Compute the marginal distribution of X_1.
   (d) Compute the predictive distribution of $X_{n+1}|X_1, \ldots, X_n$. Of course, we now assume that $X_1, \ldots, X_n, X_{n+1}|\theta$ are conditionally i.i.d.

   Hint: The densities p_θ are in an exponential family and the prior is conjugate.