Statistics & Data Sciences: First Year Prelim Exam
May 2018

Instructions:

1. Do not turn this page until instructed to do so.

2. Start each new question on a new sheet of paper.

3. This is a closed book exam.

4. Show your working and explain your reasoning as fully as possible.

5. Switch off all electronic devices.

6. The time limit is 3 hours.

7. The exam has 6 questions – do as much as you can but credit will be given for only 5 answers.
Question 1. Consider the exponential family model given by

\[ f_Y(y; \theta) = c(y) \exp \{ y\theta - b(\theta) \} \]

where \( c \) is a known function and \( b \) is part of the normalizing constant.

(i) Show that \( \mathbb{E} Y = b'(\theta) \), where \( b' \) denotes the derivative of \( b \), assumed to exist.

(ii) Prove that \( b \) is a convex function; i.e. \( b''(\theta) \geq 0 \), assumed to exist.

(iii) Show that \( \text{Var} Y = b''(\theta) \).

The MLE \( \hat{\theta} \) is given by

\[ b'(\hat{\theta}) = \bar{Y}, \]

where \( \bar{Y} \) is the sample mean.

(iv) Show that if the true model is \( f_0(y) \); i.e. the data arise from \( f_0 \), show that \( \hat{\theta} \) converges to the \( \theta^* \) which maximizes

\[ \int f_0(y) \log f(y; \theta) \, dy. \]
Question 2. Suppose $X_1, \ldots, X_n$ are independent and identically distributed from some distribution function $F(x)$ and the aim is to test

$$H_0 : F(0) = \frac{1}{2} \quad \text{vs} \quad H_1 : F(0) \neq \frac{1}{2}.$$ 

That is, the test is to see whether the median of $F$ is at 0 or not.

(i) If $H_0$ is true, what is the distribution of

$$T = \sum_{i=1}^{n} 1(X_i \leq 0).$$

Here $1(X \leq 0) = 1$ if $X \leq 0$ and otherwise is 0.

(ii) Hence, using a normal approximation for $T$, which will involve finding the mean and variance of $T$, describe how you would test the hypothesis; i.e. find the test statistic and how to obtain the critical value.

(iii) Find the Laplace transform of a normal density function with mean $\mu$ and variance $\sigma^2$; i.e. Find

$$E e^{\theta X}$$

where $X \sim N(\mu, \sigma^2)$.

(iv) Using Laplace transforms, show that

$$\frac{T/n - F(0)}{\sqrt{F(0)(1 - F(0))/n}}$$

converges in distribution to a standard normal random variable.
Question 3. Let $P = (p_{ij})$ be a $M \times M$ transition matrix for a discrete Markov chain $(X_n)$ on \{1, \ldots, M\}, for $M > 1$. Hence, for all $n \geq 1$,

$$P(X_n = j|X_{n-1} = i) = p_{ij}.$$  

(i) Show that all the eigenvalues of $P$ lie between $-1$ and $+1$ and that $1$ is actually the largest eigenvalue.

Hint: For any vector $v' = (v_1, \ldots, v_M)$ of dimension $M$ it is that

$$v(1) \leq \sum_{j=1}^{M} p_{ij} v_j \leq v(M)$$

where $v(1)$ is the minimum from $(v_1, \ldots, v_M)$ and $v(M)$ is the maximum.

(ii) If $P$ is reducible, explain why the second largest eigenvalue of $P$ is 1.

Now suppose $\pi$ is the unique stationary probability for $P$, i.e. $\pi' P = \pi'$, and that $P$ is irreducible and that the smallest eigenvalue of $P$ is greater than $-1$. Let $(\pi, w_2, \ldots, w_M)$ be the left eigenvectors of $P$.

(iii) Prove that $w'_j \mathbf{1} = 0$ where $\mathbf{1}$ is the $M$–vector of 1’s.

(iv) The chain is started at $q_0$; i.e. $P(X_0 = j) = q_{0j}$. Show that we can write

$$q_0 = \alpha \pi + \sum_{j=2}^{M} \alpha_j w_j$$

and, using (iii), show that $\alpha = 1$.

Show that the probability mass function for $X_n$ is

$$q_n = \pi + \sum_{j=2}^{M} \alpha_j \lambda_j^n w_j$$

and hence explain why $q_n \to \pi$. 


Question 4. Assume independent observations \((y_i)_{i=1}^n\) arise from the model

\[ y_i = \theta_i + \varepsilon_i, \]

where the \(\theta = (\theta_1, \ldots, \theta_n)\) are unknown parameters and \(E\varepsilon_i = 0\) and \(Var\varepsilon_i = \sigma^2\), where \(\sigma > 0\) is assumed known. The true value for the parameter is denoted by \(\theta^*\).

(i) Find the least squares estimator \(\hat{\theta}\) of \(\theta\) subject to the constraint

\[ \sum_{i=1}^n \theta_i = \tau, \]

for some fixed \(\tau\).

(ii) Find \(E\hat{\theta}\) and \(Cov\hat{\theta}\).

Now consider the model

\[ y_i = x_i' \theta + \varepsilon_i \]

where the \(\varepsilon_i\) are the same as before, but know \(\theta\) is an unknown parameter of dimension \(d\) and \(x_i' = (x_{i1}, \ldots, x_{id})\) is a vector of covariates.

(iii) Find the least squares estimator of \(\theta\) subject to constraint

\[ \theta'\theta = \tau \]

for some fixed \(\tau > 0\).

(iv) Show that the estimator given in (iii) has a Bayesian interpretation.
**Question 5.** Consider the model for which

\[
P(y_i = k | x_i, \theta, \beta) = \frac{\left( \theta e^{\beta x_i} \right)^k}{k!} \exp \left\{ -\theta e^{\beta x_i} \right\}
\]

for \( k \in \{0, 1, 2, \ldots\} \). So this is a Poisson regression model. The prior for \( \theta \) is \( \text{gamma}(a, b) \) and the prior for \( \beta \) is \( \text{N}(0, \sigma^2) \), where \((\sigma, a, b)\) is specified.

(i) Write down the posterior distribution for \((\theta, \beta)\).

(ii) Describe in detail a Gibbs sampler for approximately sampling from the posterior. In particular, show that the transition density for your chain has the posterior as the stationary density.

(iii) If the above model refers to model \( M_1 \), then \( M_2 \) is the same model except now \( \beta = 0 \) is fixed. If the prior for each model is \( 1/2 \); i.e. \( P(M_1) = P(M_2) = 1/2 \), give a method by which you could compute \( P(M_1 | \text{data}) \).

(iv) Give a full interpretation of the probabilities for the models; given that if model \( M_2 \) is correct then so is model \( M_1 \). This interpretation should include the cases when one of the models is correct and when neither is correct.
Question 6. Suppose \( p_\theta(x) \) is a density function for each \( \theta \in \Theta \) and \( p_\theta(x) \) a member of the family for a particular \( \theta_0 \) and observations \((x_1, \ldots, x_n)\) are i.i.d. from \( p_\theta \). The aim is to estimate \( \theta \) by minimizing the Fisher information distance given by

\[
F(p_0, p_\theta) = \int p_0(x) \left( \frac{p_\theta'(x)}{p_\theta(x)} - \frac{p_\theta'(x)}{p_0(x)} \right)^2 \, dx,
\]

where \( p_\theta'(x) \) denotes the derivative with respect to \( x \).

(i) Show that under regularity conditions, we can write

\[
F(p_0, p_\theta) = \kappa + \int p_0(x) \left\{ \left( \frac{p_\theta'(x)}{p_\theta(x)} \right)^2 + 2 \frac{d}{dx} \left( \frac{p_\theta'(x)}{p_\theta(x)} \right) \right\} \, dx,
\]

where \( \kappa \) is a constant not depending on \( \theta \).

(ii) If

\[
p_\theta(x) = \frac{\exp(w(x)/\theta)}{\int \exp(w(x)/\theta) \, dx}
\]

find

\[
l(\theta, x) = \left( \frac{p_\theta'(x)}{p_\theta(x)} \right)^2 + 2 \frac{d}{dx} \left( \frac{p_\theta'(x)}{p_\theta(x)} \right).
\]

(iii) Hence, explain how one could estimate \( \theta \) from the observations using an approximation to \( F(p_0, p_\theta) \).

(iv) Find the form of the estimator using your answer to (iii) with the model in (ii).