BASKIN SCHOOL OF ENGINEERING

Department of Applied Mathematics and Statistics

2014 First Year Exam: June 16, 2014

INSTRUCTIONS

If you are on the Applied Mathematics track, you are required to complete problems 1(AMS 203), 2(AMS 211), 3(AMS 212A), 4(AMS 212B), 5(AMS 213), and 6(AMS 214).

If you are on the Statistics track, you are required to complete problems 1(AMS 203), 2(AMS 211), 7(AMS 205B), 8(AMS 206B), 9(AMS 207), and 10(AMS256).

Please complete all required problems on the **supplied exam papers**. Write your exam ID number and problem number on each page. Use only the **front side** of each page.

Problem 1 (AMS 203):

Let X_1, X_2, X_3 be independent and identically distributed exponential random variables with density

$$f_X(x) = \begin{cases} \beta e^{-\beta x} & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Let $Y_i = X_1 + \ldots + X_i$, for $i = 1 : 3$.

1. (30%) Find the joint p.d.f. of Y_1 and Y_2.
2. (20%) Find the p.d.f. of Y_2.
3. (30%) Find the joint p.d.f. of Y_1, Y_2 and Y_3.
4. (20%) Show that the p.d.f. of Y_3 is

$$f_{Y_3}(y_3) = \begin{cases} \beta^3 \frac{y_3^2}{2} e^{-\beta y_3} & \text{for } y_3 > 0, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, show that Y_3 is a gamma random variable with parameters 3 and β.

Problem 2 (AMS 211):

Problem 3 (AMS 212A):

Problem 4 (AMS 212B):

Problem 5 (AMS 213):

Problem 6 (AMS 214):

Problem 7 (AMS 205B):

This problem concerns some things that can go wrong with frequentist inference when Your sample size is small.

You have a single observation Y, which You know is drawn from the Beta$(\theta, 1)$ distribution, but $\theta \in (0, 1)$ is unknown to You; Your goal here is to construct good point and interval estimates of θ. Recall that in this sampling model (a) $p(y|\theta) = \theta\, y^{\theta-1}\, I(0 < y < 1)$, where $I(A) = 1$ if proposition A is true and 0 otherwise, and (b) the repeated-sampling mean of Y is $E_{RS}(Y|\theta) = \frac{\theta}{\theta+1}$.

(1) (15%) Consider *any* estimator $\hat{\theta}(y)$ of θ in this model, where y is a realized value of Y. Identify all of the following qualitative behaviors that would be desirable for $\hat{\theta}(y)$ here, explaining briefly in each case, or explain briefly why none of them would be desirable (if none are). *Hint:* You may find it helpful to sketch the likelihood function (arising from this sampling model) for several values of y.

   (a) $\lim_{y \downarrow 0} \hat{\theta}(y) = 0$.

   (b) $\hat{\theta}(y)$ should be a non-decreasing function of y.

   (c) $\lim_{y \uparrow 1} \hat{\theta}(y) = 1$.

(2) (15%) Find the method-of-moments estimate $\hat{\theta}_{MoM}$ of θ, and identify the complete set of observed values y of Y under which $\hat{\theta}_{MoM}$ fails to respect the range restrictions on θ. Which, if any, of the desirable qualitative behaviors in (1) does $\hat{\theta}_{MoM}$ exhibit? Explain briefly.

(3) (15%) Repeat all parts of (2) for the maximum-likelihood estimate $\hat{\theta}_{MLE}$, showing that in this problem $\hat{\theta}_{MLE} = \min\left(-\frac{1}{\log Y}, 1\right)$. Does this approach exhibit better qualitative performance than $\hat{\theta}_{MoM}$ here? Explain briefly.

(4) (20%) Use observed information to construct an approximate 95% MLE-based confidence interval for θ, identifying the assumptions built into this method and commenting on whether they're likely to be met in this case. Are there values of y for which this confidence interval can violate the range restrictions on θ? If so, identify the relevant set of y values; if not, briefly explain why not.

(5) (35%) Show that in this model $V = \frac{\theta}{-\frac{1}{\log Y}} = -\theta \log Y$ is a pivotal quantity, in this case with the standard Exponential distribution $p(v) = e^{-v}$, and use this to construct (for any given $\alpha \in (0,1)$) an exact $100(1-\alpha)\%$ confidence interval for θ based on the single observation Y. As a function of Y, can anything go wrong with this interval with respect to range restrictions on θ? Explain briefly.

Problem 8 (AMS 206B):

Assume a Bernoulli experiment in which you perform n independent trials with success probability θ, and X counts the number of successes. Then, $X|\theta \sim \text{Binomial}(n, \theta)$ with n known.

1. (20%) Assume you observe $X = x$. Find the posterior density of θ, $\pi(\theta|x)$, under a uniform prior $\theta \sim U(0, 1)$.

2. (40%) Consider the loss function defined by

$$L(\theta, \hat{\theta}) = \frac{(\hat{\theta} - \theta)^2}{\theta(1 - \theta)}.$$

   Find $\hat{\theta}(x)$ the estimator that minimizes the Bayesian expected posterior loss under the scenario described in part 1. (Note: Assume $x \neq 0$.)

3. (40%) Find Jeffreys prior for θ and the corresponding posterior distribution under such prior.

Some information that may be useful:

- If $X \sim \text{Binomial}(n, \theta)$, X has probability mass function given by

$$\Pr(X = x) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}, \quad x = 0, 1, \ldots, n.$$

- Beta distribution. If $X \sim \text{Beta}(\alpha, \beta)$ its pdf is given by

$$p(x|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1}, \quad 0 \leq x \leq 1.$$

  In addition, $E(X) = \alpha/(\alpha + \beta)$, and $V(X) = \alpha\beta/\{(\alpha + \beta)^2(\alpha + \beta + 1)\}$.

- $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$.

Problem 9 (AMS 207):

Consider the linear model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \;\; \boldsymbol{\varepsilon} \sim N_n(\boldsymbol{0}, (1/\phi)\boldsymbol{I})$$

where $\boldsymbol{y} \in \mathbb{R}^n$, $\boldsymbol{X}$ is a full rank matrix of dimensions $n \times p$, $\boldsymbol{\beta} \in \mathbb{R}^p$, $\boldsymbol{I}$ is the $n \times n$ identity matrix, and $\phi > 0$. Consider the prior $p(\boldsymbol{\beta}|\phi) = \prod_i p(\beta_i|\phi)$, with

$$p(\beta_i|\phi) = \left(1 + \frac{\phi\beta_i^2}{\nu\lambda}\right)^{-(\nu+1)/2} \frac{\phi}{\lambda} \;\; ,$$

where ν and λ are known, and assume that $p(\phi) \propto 1/\phi$.

1. (30%) Write $p(\beta_i|\phi)$ as a scale mixture of normals.
2. (20%) Use the above representation to introduce latent variables that facilitate sampling the posterior distribution of all parameters using Gibbs sampling. Write the resulting model.
3. (50%) Obtain the full conditionals for all model parameters.

Some information that may be useful:

- Assuming the matrices and vectors below are of appropriate dimensions (and also that matrix $A + B$ is invertible), we have

  $$(x-a)'A(x-a)+(x-b)'B(x-b) = (x-c)'(A+B)(x-c)+(a-b)'A(A+B)^{-1}B(a-b)$$

  where

  $$c = (A + B)^{-1}(Aa + Bb)$$

- $$\int_0^\infty \frac{e^{-\beta/x}}{x^{\alpha+1}} \frac{\beta^\alpha}{\Gamma(\alpha)} dx = 1$$

Problem 10 (AMS 256):

Consider the random blocks model $y_{i,j} = \mu + \alpha_i + \beta_j + \epsilon_{i,j}$, with $\epsilon_{i,j} \sim \mathsf{N}(0, \sigma^2)$, for $i = 1, 2, 3$ and $j = 1, 2, 3$.

1. (15%) Write this model in matrix form as $\mathbf{y} = \mathbf{X}\boldsymbol{\theta}+\boldsymbol{\epsilon}$, where $\boldsymbol{\theta} = (\mu, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3)$.

2. (10%) What is the rank of the design matrix $\mathbf{X}$? Justify your answer.

3. (35%) List one possible set of constraints on $\boldsymbol{\theta}$ that makes the corresponding constrained least squares estimator be unique, and provide the solution to the normal equations under those constraints. (Hint: Rather than solving the normal equations, try to propose a solution and show that it satisfies the normal equations.)

4. Suppose that we are interested in testing the hypotheses $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$ against the alternative H_a : at least one $\alpha_i \neq 0$.

   (a) (10%) Write these hypotheses as a general linear hypotheses $H_0 : \mathbf{K}^T\boldsymbol{\theta} = \mathbf{m}$ and $H_a : \mathbf{K}^T\boldsymbol{\theta} \neq \mathbf{m}$. Make sure to show that $\mathbf{K}^T\boldsymbol{\theta} = \mathbf{m}$ is testable!

   (b) (30%) Describe a test (i.e., a statistic, its distribution under the null and the alternative, and a rejection region) for the general linear hypotheses described before.