BASKIN SCHOOL OF ENGINEERING

Department of Applied Mathematics and Statistics

2014 First Year Exam Retake: September 17, 2014

INSTRUCTIONS

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):

Consider an electronic system comprised of three components denoted as C_1, C_2, and C_3. Let X_i be a random variable denoting the lifetime of component C_i, for $i = 1, 2, 3$. Assume X_1, X_2, and X_3 are i.i.d. random variables each with p.d.f.

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

Suppose that the system will operate as long as both component C_1 and at least one of the components C_2 and C_3 operate. Let Z be the random variable that denotes the lifetime of the system.

1. (60%) Find the c.d.f. of Z.
2. (40%) Obtain $\mathrm{E}(Z^2)$.

Useful information: A r.v. X follows an exponential distribution with parameter $\beta > 0$ if its p.d.f. is given by

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

In this case $\mathrm{E}(X) = 1/\beta$ and $\mathrm{Var}(X) = 1/\beta^2$.

Problem 2 (AMS 211):

1. (50%) Solve
$$x\frac{dy}{dx} - 2y = x^4 \sin x$$

2. (50%) Find the eigenvalues and eigenvectors of
$$\begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$$

Problem 3 (AMS 205B):

For each statement below, say whether it's true or false; if true without further assumptions, briefly explain why it's true and describe its implications for statistical inference; if it's sometimes true, give the extra conditions necessary to make it true; if it's false, briefly explain how to change it so that it's true and/or give an example of why it's false.

1. (25%) When the set Θ of possible values of a (one-dimensional) real-valued parameter θ has one or more boundaries beyond which θ doesn't make scientific sense, an unbiased estimator $\hat{\theta}_U$ of θ can misbehave (in the sense of taking values that are outside Θ). However, when Θ is the entire real line, this sort of absurd behavior of $\hat{\theta}_U$ cannot occur, and $\hat{\theta}_U$ may provide a reasonably good estimate of θ.

2. (25%) Consider the sampling model $(Y_i|\eta, \mathcal{B}) \stackrel{\text{IID}}{\sim} p(y_i|\eta, \mathcal{B})$, for $i = 1, \ldots, n$, where the Y_i are univariate real values, η is a parameter vector of length $1 \leq k < \infty$, and $\mathcal{B}$ summarizes Your background information; a Bayesian analysis with the same sampling model would add a prior distribution layer of the form $(\eta|\mathcal{B}) \sim p(\eta|\mathcal{B})$ to the hierarchy. The Bernstein-von Mises theorem says that maximum-likelihood (ML) and Bayesian inferential conclusions will be similar in this setting if (a) n is large and (b) $p(\eta|\mathcal{B})$ is diffuse, but the theorem does not provide guidance on how large n needs to be for its conclusion to hold in any specific sampling model.

3. (25%) Method-of-moments (MoM) estimators are as efficient as ML estimators, but MoM estimators have the disadvantage of often not being expressible in closed form, whereas ML estimators always have closed-form algebraic expressions.

4. (25%) When Your sampling model has n observations and a single parameter θ (so that $k = 1$ in the notation of part 2), if the sampling model is regular (i.e., if the range of possible data values doesn't depend on θ), in large samples the observed information $\hat{I}\left(\hat{\theta}_{MLE}\right)$ is $O(n)$, meaning that (a) information in $\hat{\theta}_{MLE}$ about θ increases linearly with n and (b) $\hat{V}\left(\hat{\theta}_{MLE}\right) = O\left(\frac{1}{n}\right)$.

Problem 4 (AMS 206B):

Consider the estimation of the parameter $\theta \in (0, \infty)$ under the loss function

$$L(\theta, d) = \frac{(\theta - d)^2}{\theta(\theta+1)}, \qquad (1)$$

based on one observation $X = x$ from the negative binomial distribution parameterized as

$$f(x|\theta) = \binom{n+x-1}{x} \theta^x (\theta+1)^{-(n+x)}, \quad x \in \{0, 1, 2, ...\}$$

where n is known. Note that, under this parameterization, $\mathrm{E}(X|\theta) = n\theta$ and $\mathrm{Var}(X|\theta) = n\theta(\theta+1)$.

1. (20%) Determine the risk function of the unbiased estimator $\delta_0(x) = x/n$.
2. (20%) Determine the risk function of the estimator $\delta_1(x) = x/(n+1)$.
3. (10%) Which of the two estimators ($\delta_0(x)$ or $\delta_1(x)$) has lower maximum risk? Justify your answer.
4. (15%) Consider the following family of prior distributions for θ,
$$\pi(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (\theta+1)^{-(a+b)}, \quad a > 0,\ b > 0. \qquad (2)$$
Obtain the posterior distribution for θ under the prior in (2).
5. (35%) Find the Bayes rule (estimator) under the loss function in (1) and the prior given in (2).

Some facts that may be useful:

- Under the squared error loss function $L(\theta, d) = (\theta - d)^2$ we have that
$$R(\theta, \delta(x)) = \mathrm{Bias}^2(\delta(x)) + \mathrm{Variance}(\delta(x)),$$
with $R(\theta, \delta(x)) = \int L(\theta, \delta(x)) f(x|\theta) \mathrm{d}x$.
- For the prior distribution in (2), we have that
$$\mathrm{E}(\theta^c (\theta+1)^{-d}) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+c)\Gamma(b+d-c)}{\Gamma(a+b+d)}.$$

Problem 5 (AMS 207):

Consider a set of observations $x_1, \dots, x_n, x^*_{n+1}, \dots, x^*_{n+k}$. Ignore for now the asterisks, and assume that they are independently distributed according to $Exp(\lambda_i)$, for $i = 1, \dots, n+k$, where λ_i^{-1} is the mean of the exponential distribution for the ith observation. Consider the prior

$$p(\lambda_1, \dots, \lambda_{n+k}|\beta) = \prod_{i=1}^{n+k} p(\lambda_i|\beta) = \prod_{i=1}^{n+k} Ga(\lambda_i|\alpha, \beta)$$

where α is a known positive constant, and $p(\beta) = Ga(\beta|a, b)$, with a and b are both known. Here, $Ga(x|A, B)$ denotes the density of a Gamma distribution with shape parameter $A > 0$ and rate parameter $B > 0$, that is, $\{B^A x^{A-1} \exp(-Bx)\}/\Gamma(A)$, for $x > 0$. (Recall that the exponential density arises as the special case with $A = 1$.)

1. (30%) Find the unconditional prior density $p(\lambda_1, \dots, \lambda_{n+k})$. Are the λ_i independent? Are the λ_i exchangeable?

2. (25%) Find the marginal posterior distribution of β given the sample.

Next, suppose that the observations with an asterisk are right censored. That is, the true value is missing, but we know that it is larger than the value recorded with an asterisk.

3. (20%) Use the missing value approach to obtain the likelihood that accounts for the censoring.

4. (25%) Find the marginal posterior distribution of β in this case. Is there any difference with respect to the case where the censoring is ignored?

Problem 6 (AMS 256):

Consider the model $y_i = \mu + \beta \sin(x_i + \eta) + \epsilon_i$, where $x_i = \frac{2\pi i}{n}$ for $i = 0, 1, \ldots, n$, $\epsilon_i \sim \mathsf{N}(0, 1)$ independently for each i, and $\eta \in [0, \pi]$. Here, μ, β, and η are the unknown model parameters.

1. (50%) Use the trigonometric identity, $\beta \sin(x_i + \eta) = \beta \cos(\eta) \sin(x_i) + \beta \sin(\eta) \cos(x_i)$, to write the model using a linear model formulation with a three-dimensional vector of regression coefficients. Based on the linear model formulation, obtain the maximum likelihood estimator for (μ, β, η).
   (**Hint:** Note that $\sum_{i=0}^{n} \sin(x_i) = 0$, and $\sum_{i=0}^{n} \sin(x_i) \cos(x_i) = 0$.)

2. (50%) Assume that you are interested in testing the hypotheses $H_0 : \eta = \pi/4$ vs. $H_a : \eta \neq \pi/4$. Derive the likelihood ratio test for this pair of hypotheses. (Please keep in mind that we are assuming that the error variance is known!)