— 2012

BASKIN SCHOOL OF ENGINEERING

Department of Applied Mathematics and Statistics

First Year Exam: June 18th, 2012

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 and X_2 be two independent random variables each with p.d.f. $f_X(x) = e^{-x}, x > 0$ and $f_X(x) = 0, x \leq 0$. Let $Z = X_1 - X_2$ and $W = X_1/X_2$

1. (25 points) Show that the joint p.d.f. of $Y = X_1$ and Z is given by:

$$g(y, z) = \exp(-2y + z)$$

and find the region where this p.d.f. is defined.

2. (15 points) Prove that the conditional p.d.f. of X_1 given $Z = 0$ is:

$$h(x_1|0) = \begin{cases} 2e^{-2x_1} & \text{for } x_1 > 0 \\ 0 & \text{otherwise} \end{cases}$$

3. (25 points) Show that the joint p.d.f. of $Y = X_1$ and W is given by:

$$g(y, w) = y\exp(-y(1 + 1/w))/w^2$$

and find the region where this p.d.f. is defined.

4. (15 points) Prove that the conditional p.d.f. of X_1 given $W = 1$ is:

$$h(x_1|1) = \begin{cases} 4x_1e^{-2x_1} & \text{for } x_1 > 0 \\ 0 & \text{otherwise} \end{cases}$$

5. (20 points) Note that $\{Z = 0\} = \{W = 1\}$, but the conditional distribution of X_1 given $Z = 0$ is not the same as the conditional distribution of X_1 given $W = 1$. Can you comment on the results in Parts 2 and 4?.

Problem 2 (AMS 211):

1. [30%] Calculate the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix}$$

2. [40%] A function $f(x)$ equals $1 - x^2$ over $0 < x < 1$. Before doing any calculations, sketch the three possible periodic extensions of the nonperiodic function in the doubled interval that represent the three different possible Fourier series expansions of the function. Briefly state which one you expect to work the best and why. Calculate the Fourier coefficients for that series.

3. [30%] Solve

$$\frac{d^2y}{dx^2} - 7\frac{dy}{dx} + 10y = e^{3x}$$

subject to boundary conditions

$$y(0) = -\frac{1}{2}, \quad \frac{dy}{dx}(0) = \frac{5}{2}$$

Problem 3 (AMS 212A):

Show (i.e. do not merely verify), making sure to write *all* steps of the process, that the steady state temperature distribution $T(r,\theta,z)$ in a solid semi-infinite thermally conducting cylindrical rod, satisfying the following equation and boundary conditions:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 T}{\partial \theta^2}+\frac{\partial^2 T}{\partial z^2}=0 \text{ for } r<1\ ,\ z>0$$
$$T(1,\theta,z)=0$$
$$T(r,\theta,0)=r\sin\theta$$
$$T\to 0 \text{ as } z\to+\infty$$

is

$$T(r,\theta,z)=\sum_{n=1}^{\infty}\frac{2}{a_n J_2(a_n)}J_1(a_n r)e^{-a_n z}\sin(\theta)$$

where a_n is the n-th zero of the J_1 Bessel function.

Hints:

- The equation $x^2f''+xf'+(x^2-n^2)f=0$ is a Bessel function, and has regular solutions $J_n(x)$ and singular solutions $Y_n(x)$.
- Given that the boundary conditions only depend on $\sin\theta$, you can save time by considering only terms that involve the function $J_1(x)$. Briefly explain why.
- The orthogonality condition for the Bessel functions is

$$\int_0^1 xJ_n(ax)J_n(bx)dx=0 \text{ if } a\neq b$$
$$\int_0^1 xJ_n^2(ax)dx=\frac{1}{2}J_{n+1}^2(a)$$

for any n as long as a and b are zeros of $J_n(x)$.

- You will need to use the fact that

$$\int_0^1 x^2J_1(ax)dx=\frac{J_2(a)}{a}$$

for any a that is a zero of $J_1(x)$.

Problem 4 (AMS 212B):

1. (25%) Consider the Dawson function defined as

$$F(x)\equiv\exp(-x^2)\int_0^x\exp(t^2)dt$$

Find the first 2 terms in the expansion of $F(x)$ as $x\to+\infty$.

2. (75%) The horizontal oscillation of a mass attached to the middle point of a stretched vertical spring is described by

$$\begin{cases} y''+\left(2-\dfrac{1}{\sqrt{1+\varepsilon^2y^2}}\right)y=0 \\ y(0)=1\ ,\quad y'(0)=0\end{cases} ,\qquad \varepsilon\to 0_+$$

where ε is the nondimensionalized amplitude of oscillation. Use the method of strained variable to solve the initial value problem.

- Find the first two terms (up to ε^2 term) in the expansion.
- Find the period of oscillation (up to ε^2 term). Does the period increase or decrease when the oscillation amplitude is increased?

Hint: You may need

$$\cos^3(s)=\frac{1}{4}\left(\cos(3s)+3\cos(s)\right)$$

If you use the Laplace transform in solving ODE's (you don't have to use the Laplace transform), you may need

$$L[f''(t)]=s^2L[f(t)]-sf(0)-f'(0)$$
$$L[\cos(\alpha t)]=\frac{s}{s^2+\alpha^2}$$

Problem 5 (AMS 213):

Part 1 [50%] Consider the least square problem of minimizing

$$r(x)=(Ax-b)^T(Ax-b),$$

where A is an $m \times n$ matrix with $m > n$. It is assumed that A has full rank. Let the full QR factorization of A be

$$A = [Q_1, Q_2]\begin{bmatrix} R \\ 0 \end{bmatrix},$$

where R is an $n \times n$ upper triangular matrix, Q_1 is $m \times n$, and Q_2 is $m \times (m-n)$. Let x^* be the solution of the least square problem.

1. Show that x^* satisfies equation $Rx = Q_1^T b$. [30%]
2. Show that $r(x^*) = \|Q_2^T b\|_2^2$. [20%]

Part 2 [50%] Consider the following numerical scheme

$$x_{k+1} = x_k + hf(0.5x_k + 0.5x_{k+1}) \qquad (1)$$

for solving the initial value problem of $dx/dt = f(x)$.

1. Is scheme (1) explicit or implicit, single step or multistep?[10%]
2. Show that scheme (1) is Absolutely Stable. [40%]

<u>Problem 6 (AMS 214):</u>

Consider the system

$$\dot{x} = x^2 - y - 1 \qquad (2)$$
$$\dot{y} = y(x-2) \qquad (3)$$

a) (30%) Find three fixed points and classify them.

b) (30%) Find the eigenvalues and eigenvectors of the linearized dynamical system at the fixed points, and use them to sketch the phase portrait.

c) (10%) State what the index of each fixed point is.

d) (30%) Based on the sketch, you can assume that the three straight lines through pairs of fixed points contain straight-line trajectories. Use this assumption and the index theory to show that there are no closed orbits. Explain your reasoning.

<u>Problem 7 (AMS 205B):</u>

Consider estimating the number $0 < N < \infty$ of individuals in a finite population (such as $\mathcal{P}$ = {the deer living on the UCSC campus as of 1 July 2012}). One popular method for performing this estimation is *capture-recapture* sampling; the simplest version of this approach proceeds as follows. In stage I, a random sample of m_0 individuals is taken, and all of these individuals are tagged and released; then, a short time later, in stage II a second independent random sample of n_1 individuals is taken, and the number m_1 of these n_1 individuals who were previously tagged is noted.

There are a number of ways to perform the random sampling in stages I and II; the least complicated methods are IID sampling (at random with replacement) and simple random sampling (SRS; at random without replacement). For the rest of the problem, suppose that you've decide to use SRS at stage I and IID sampling at stage II, which will be denoted (SRS, IID).

(a) (10 points) Briefly explain why the following sampling model follows naturally from the scientific context of the problem under (SRS, IID):

$$(m_1|N) \sim \text{Binomial}\left(n_1, \frac{m_0}{N}\right). \qquad (4)$$

(b) (Estimation of N)

(i) (20 points) Show that in model (4) the method-of-moments estimator $\hat{N}_{MoM}$ and the maximum-likelihood estimator $\hat{N}_{MLE}$ coincide and are given by

$$\hat{N}_{MoM} = \hat{N}_{MLE} \equiv \hat{N} = \frac{n_1 m_0}{m_1} \qquad (5)$$

(give a full argument for why (5) is the global maximum of the likelihood or log-likelihood function).

(ii) (6 points) Given the nature of (SRS, IID) sampling, why is this estimator intuitively sensible? Explain briefly.

(c) (Uncertainty assessment)

(i) (17 points) Use the Δ-method to show that in this model the repeated-sampling variance of $\hat{N}$ is approximately

$$V_{RS}\left(\hat{N}\right) \doteq \frac{N^2(N-m_0)}{n_1 m_0}. \quad (6)$$

(ii) (6 points) If $m_0 \to N$ were physically possible to attain in sampling from $\mathcal{P}$ using (SRS, IID), briefly explain why it makes good scientific sense that (6) goes to 0 in this limit.

(iii) (6 points) Does it make good scientific sense that (6) also goes to 0 under a scenario in which m_0 is held fixed and $n_1 \to \infty$? Explain briefly.

(iv) (15 points) Use observed Fisher information to compute an approximation to the repeated-sampling estimated variance $\hat{V}_{RS}\left(\hat{N}\right)$ of $\hat{N}_{MLE}$, thereby showing that this approximation coincides with the obvious estimate of (6) in this case.

(v) (20 points) Use your calculation in part (c) (iv) to give an approximate 95% confidence interval for N, and explain briefly under what conditions you expect this interval to be close to frequentist-valid (and why).

Problem 8 (AMS 206B):

Let $x_1, \ldots, x_n$ be an independent and identically distributed sample such that $x_i \sim \mathsf{N}(\theta, 1)$, where θ is unknown. Also, assume that it is known that $\theta \geq 0$ and that an exponential prior with mean λ is reasonable.

1. (50 points) Compute the maximum a posteriori (MAP) estimator for θ.
2. (15 points) What utility function justifies the use of this estimator?
3. (35 points) Assuming that $\bar{x} - \frac{1}{n\lambda} < 0$, compute the 95% HPD credible interval for θ. Write your response in terms of Φ, the cumulative distribution function for the standard normal distribution, and its inverse.

Problem 9 (AMS 207):

Consider the linear model

$$y = X\beta + \epsilon, \quad \epsilon \sim N_p(0, 1/\phi V_\psi)$$

where $y \in \mathbb{R}^n$, X is a full rank matrix of dimensions $n \times p$, $\beta \in \mathbb{R}^p$ and V_ψ is a $p \times p$ matrix that is a known function of ψ.

1. (35 points) Assume that ψ is known. Find the sufficient statistics for β and ϕ.
2. (45 points) Consider the prior

$$p(\beta, \phi|\psi) \propto \frac{1}{\phi} N_p(0, g/\phi(X'V_\psi^{-1}X)^{-1})$$

Write the posterior distribution of β and ϕ as

$$p(\beta, \phi|y, X, \psi) = p(\beta|\phi, y, X, \psi)p(\phi|y, X, \psi)$$

and specify the densities for each one of the two factors in this expression as a function of the sufficient statistics.

3. (20 points) Suppose ψ is unknown. Obtain the posterior $p(\psi|y, X)$.

Hint: Recall that

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

Problem 10 (AMS 256):

Suppose that we have the following data

Response y_{ij}		Covariate x_{ij}	
Type I	Type II	Type I	Type II
4	2	2	3
6	10	1	2
17	21	-3	-5

Here x_{ij} is a covariate that is measured at the same time as the response for each experimental unit ($j = 1, 2, 3$) in each type ($i = 1, 2$).

Consider the following model:

$$y_{ij} = \mu + \alpha_i + \beta x_{ij} + e_{ij}, \quad \epsilon_{ij}, \tag{7}$$

with the ϵ_{ij}s i.i.d. $N(0, \sigma^2)$.

1. **PART I.**

   (a) (15 points) Write down the model in matrix form $\mathbf{y} = \mathbf{X}\beta + \epsilon$. What is the rank of $\mathbf{X}$? What is the rank of $\mathbf{X}'\mathbf{X}$?

   (b) (10 points) Let $\hat{\beta}$ be a solution to the normal equations $\mathbf{X}'\mathbf{X}\beta = \mathbf{X}'\mathbf{y}$. Is $\hat{\beta}$ unique? Let $\hat{y} = \mathbf{X}\hat{\beta}$. Is $\hat{y}$ unique? Justify your answers.

   (c) (8 points) Is $H_0: \ \alpha_1 = 0$ testable? Justify your answer.

   (d) (9 points) Is $H_0: \ \alpha_1 - \alpha_2 = 0$ testable? Is the B.L.U.E. of $\alpha_1 - \alpha_2$ unique? Justify your answers.

   (e) (8 points) Is $H_0: \ \beta = 0$ testable? Justify your answer.

2. **PART II.** Add the following restriction to the model in (1): $\alpha_2 = 0$. For the following questions please provide numerical values based on the data in addition to general mathematical expressions.

   (a) (20 points) Find the LSEs of μ, α_1 and β in the restricted model.

   (b) (30 points) Let $H_0: \alpha_1 = 0$. Is H_0 testable in the restricted model? Is so, derive the F-statistic for testing H_0.