Department of Applied Mathematics and Statistics
The Johns Hopkins University

INTRODUCTORY EXAMINATION–FALL SESSION

August 27, 2008

Instructions: Read carefully!

1. This **closed-book** examination consists of 15 problems, each worth 5 points. The passing grade has been set at 50 points, i.e. 2/3 of the total points. Partial credit will be given as appropriate; each part of a problem will be given the same weight. If you are unable to prove a result asserted in one part of a problem, you may still use that result to help in answering a later part.

2. You have been provided with a syllabus indicating the scope of the exam. Our purpose is to test not only your knowledge, but also your ability to apply that knowledge, and to provide mathematical arguments presented in **clear, logically justified steps**. The grading will reflect that broader purpose.

3. The problems have not been grouped by topic, but there are roughly equally many mainly motivated by each of the three areas identified in the syllabus (linear algebra; real analysis; probability;). Nor have the problems been arranged systematically by difficulty. If a problem directs you to use a particular method of analysis, you *must* use it in order to receive substantial credit.

4. Start your answer to each problem on a NEW sheet of paper. Write only on ONE SIDE of each sheet, and please do not write very near the margins on any sheet. Arrange the sheets in order, and write your NAME and the PROBLEM NUMBER on each sheet.

5. The examination will begin at 8:30 AM; lunch and refreshments will be provided. The exam will end just before 5:00 PM. You may leave before then, but in that case you may not return.

6. Paper will be provided, but you should bring and use writing instruments that yield marks dark enough to be read easily.

7. **No calculators of any sort are needed or permitted.**

1. The Legendre transform of a convex function $F(\mathbf{x})$ on $\mathbb{R}^n$ is another convex function on $\mathbb{R}^n$ defined by

$$F^*(\mathbf{y}) = \sup_{\mathbf{x}}[(\mathbf{x},\mathbf{y}) - F(\mathbf{x})],$$

where $(\mathbf{x},\mathbf{y})$ is the standard Euclidean inner product. If $Q_A(\mathbf{x}) = \frac{1}{2}(\mathbf{x},\mathbf{Ax})$ is a quadratic form defined by a non-singular real, symmetric $n \times n$ matrix $\mathbf{A}$, calculate $Q_A^*(\mathbf{y})$.

Solution: By multivariable calculus, the supremum is achieved at the vector $\mathbf{x}$ which satisfies

$$\mathbf{0} = \frac{\partial}{\partial \mathbf{x}}[(\mathbf{x},\mathbf{y}) - \frac{1}{2}(\mathbf{x},\mathbf{Ax})] = \mathbf{y} - \mathbf{Ax},$$

or $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$. It follows that

$$Q_A^*(\mathbf{y}) = (\mathbf{A}^{-1}\mathbf{y},\mathbf{y}) - \frac{1}{2}(\mathbf{A}^{-1}\mathbf{y},\mathbf{A}\cdot\mathbf{A}^{-1}\mathbf{y}) = \frac{1}{2}(\mathbf{y},\mathbf{A}^{-1}\mathbf{y}).$$

---

2. Find the inverse of the $n \times n$ matrix $A := J - I$, where J is the matrix with every entry equal to 1 and I is the identity matrix.

Solution: Solution #1:

Perhaps the student remembers that the inverse is $\frac{1}{n-1}J - I$. The proof is then straightforward verification, using $J^2 = nJ$:

$$(J-I)\left(\frac{1}{n-1}J - I\right) = \frac{n}{n-1}J - J - \frac{1}{n-1}J + I = I.$$

Solution #2:

If the student needs to derive the result, this is also straightforward. Let e denote the (column) vector of 1s. For general y, we solve $Ax = y$ for x, beginning with the calculation that $Ax = (e^Tx)e - x$. Each line below follows from preceding calculations, and the argument can be reversed:

$$\begin{aligned} Ax &= (e^Tx)e - x = y \\ x &= -y + (e^Tx)e \\ e^Tx &= -e^Ty + ne^Tx \\ e^Tx &= \frac{1}{n-1}e^Ty \\ x &= -y + \frac{1}{n-1}(e^Ty)e = \left(\frac{1}{n-1}J - I\right)y. \end{aligned}$$

Thus $A^{-1} = \frac{1}{n-1}J - I$.

---

3. Fix an integer $r \geq 1$. Let X be a random variable with the Gamma$(r, 1)$ distribution, i.e., with density

$$f_X(x) = \frac{1}{(r-1)!} e^{-x} x^{r-1}, \quad x > 0.$$

   (a) Calculate the moment generating function M of X. For which values of $t \in \mathbb{R}$ do we have $M(t) < \infty$?

   (b) For a real number $a > r$, derive the best Chernoff bound on the tail probability $P\{X \geq a\}$. [REMINDER: "Chernoff bound" refers to a bound on $P\{e^{tX} \geq e^{ta}\}$ for some $t > 0$ obtained using Markov's inequality.]

Solution:

(a) For $-\infty < t < 1$ we have, by the change of variables $y = (1-t)x$,

$$\begin{aligned} M(t) &= Ee^{tX} = \int_0^\infty \frac{1}{(r-1)!} e^{tx} e^{-x} x^{r-1}\, dx = (1-t)^{-r} \int_0^\infty \frac{1}{(r-1)!} e^{-y} y^{r-1}\, dy \\ &= (1-t)^{-r} < \infty, \end{aligned}$$

and for $t \geq 1$ we see that $M(t) = \infty$.

(b) Since X has moment generating function $M(t) = (1-t)^{-r}$ for $0 < t < 1$, we apply the Chernoff bound $P\{X \geq a\} \leq e^{-ta} M(t)$ and optimize over the choice of $0 < t < 1$. The best choice, quickly found to be $t = 1 - \frac{r}{a}$ by taking the logarithm of the bound, yields

$$P\{X \geq a\} \leq \exp[-(a-r)](a/r)^r.$$

---

4. The characteristic function of a random variable X is

$$\phi_X(t) = E(e^{itX}),\ -\infty < t < \infty.$$

   (a) Compute $\phi_X(t)$ if X has the uniform distribution on $[0, 1]$.

   (b) Identify a random variable with characteristic function

$$\phi(t) = \frac{2(1-\cos t)}{t^2}$$

   by considering two independent random variables, each uniform on $[0, 1]$.

Solution:

a)

$$\phi_X(t) = E(e^{itX}) = \int_0^1 e^{itx}dx = \frac{e^{it}-1}{it}.$$

b) Let $Z = X - Y$ where X and Y are independent, both be uniform on $[0,1]$. Then

$$\begin{aligned}
\phi_Z(t) &= E(e^{it(X-Y)}) \\
&= E(e^{itX})E(e^{-itY}) \\
&= \phi_X(t)\phi_Y(-t) \\
&= \left(\frac{e^{it}-1}{it}\right)\left(\frac{e^{-it}-1}{-it}\right) \\
&= \frac{2(1-\cos t)}{t^2}.
\end{aligned}$$

---

5. Let $\{a_1, a_2, ...\}$ be a sequence of real numbers for which

$$\lim_{n\to\infty} a_n = 0.$$

Prove that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} a_k = 0.$$

Solution:

Given $\varepsilon > 0$, choose N such that $|a_n| \le \varepsilon$ for $n \ge N$. Then for $n \ge N$,

$$|\frac{1}{n}\sum_{k=1}^{n} a_k| = |\frac{1}{n}\sum_{k=1}^{N} a_k + \frac{1}{n}\sum_{k=N+1}^{n} a_k| \le |\frac{1}{n}\sum_{k=1}^{N} a_k| + \varepsilon.$$

Now choose M sufficiently large that $|\frac{1}{n}\sum_{k=1}^{N} a_k| \le \varepsilon$ whenever $n \ge M$. Consequently,

$$|\frac{1}{n}\sum_{k=1}^{n} a_k| \le 2\varepsilon$$

for $n \ge \max\{N, M\}$.

---

6. Suppose that A is a $n \times n$ real matrix such that both A and $I + A$ are nonsingular. Show that

   (a) $(I+A)^{-1} = I - (A^{-1}+I)^{-1}$.

(b) $\text{trace}(I+A)^{-1}+\text{trace}(A^{-1}+I)^{-1}=n$.

Solution:

For (a), we have

$$\begin{aligned}(I+A)(I-(A^{-1}+I)^{-1}) &= (I+A)-(I+A)[(A^{-1}(I+A)]^{-1}\\ &= (I+A)-(I+A)(I+A)^{-1}A\\ &= I.\end{aligned}$$

this shows that $I-(A^{-1}+I)^{-1}$ is the inverse of $I+A$.

Then (b) follows directly from (a) and the fact that $\text{trace}(B+C)=\text{trace}(B)+\text{trace}(C)$, and $\text{trace}(I)=n$.

---

7. Show that for any rearrangement $b_1,\cdots,b_n$ of the positive numbers $a_1,\cdots,a_n$,

$$\frac{1}{n}\left(\frac{a_1}{b_1}+\frac{a_2}{b_2}+\cdots+\frac{a_n}{b_n}\right)\geq 1.$$

Solution:

This directly follows from the arithmetic-geometric mean inequality. Indeed,

$$\frac{1}{n}\left(\frac{a_1}{b_1}+\frac{a_2}{b_2}+\cdots+\frac{a_n}{b_n}\right)\geq (\frac{a_1}{b_1}\cdot\frac{a_2}{b_2}\cdots\cdot\frac{a_n}{b_n})^{\frac{1}{n}}=1.$$

For another solution, we can rearrange the ordering of the quotients a_i/b_i if necessary, and assume that $b_1\leq b_2\leq\ldots\leq b_n$. Now suppose for some particular pair of indices with $i<j$ we have $a_i>a_j$. Observe that $a_ib_j+a_jb_i\leq a_ib_i+a_jb_j$ since we can write

$$(a_ib_j+a_jb_i)-(a_ib_i+a_jb_j)=(a_i-a_j)(b_j-b_i)\geq 0.$$

Then when we swap the a_i terms in the sum of quotients, and we see that the effect is that we do not decrease the sum. That is

$$\frac{a_i}{b_i}+\frac{a_j}{b_j}=\frac{1}{b_ib_j}\left\{a_ib_j+a_jb_i\right\}$$

$$\geq\frac{1}{b_ib_j}\left\{a_ib_i+a_jb_j\right\}$$

$$\geq \frac{a_i}{b_i} + \frac{a_j}{b_j}$$

We can repeatedly swap numerators in this fashion until we attain $a_1 \leq a_2 \leq \ldots \leq a_n$. But that this point we have $a_i = b_i$ for $i = 1, \ldots, n$ so each quotient is 1, hence

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots + \frac{a_n}{b_n} \geq 1 + \ldots + 1 = n.$$

---

8. Consider a sequence $x^{(n)}$, $n = 1, 2, \ldots$ of elements of $\mathbb{R}^d$ with the following properties:

   (i) $\lim_{n\to\infty}(x_i^{(n)} - x_j^{(n)}) = 0$ for all $1 \leq i, j \leq d$,

   (ii) $\max_{1\leq i\leq d} x_i^{(n)}$, $n = 1, 2, \ldots$, is a nonincreasing sequence, and

   (iii) $\min_{1\leq i\leq d} x_i^{(n)}$, $n = 1, 2, \ldots$, is a nondecreasing sequence.

   (Here $x_i^{(n)}$ denotes the i-th coordinate of $x^{(n)}$.) Prove that $x_i^{(n)}$, $n = 1, 2, \ldots$ converges to the same limit for $i = 1, \ldots, d$

   Solution:

   Let $c_n = \min_{1\leq i\leq d} x_i^{(n)}$ and $d_n = \max_{1\leq i\leq d} x_i^{(n)}$. Then $\{c_n\}$ is nondecreasing and $\{d_n\}$ is nonincreasing and $c_n \leq d_n$ for all n, and $c_n \leq x_i^{(n)} \leq d_n$ for $n = 1, 2, 3, \ldots$ and $1 \leq i \leq d$. There are two possibilities to consider. Either these two sequences covnverge to the same limit, in which case it is clear that all of the $x_i^{(n)}$ converge to this limit as well.

   The other possibility is that $c_n \to c$ and $d_n \to d$ for some $c < d$. In this case, let i_n be an index for which $x_{i_n}^{(n)} = \min_{1\leq i\leq d} x_i^{(n)}$ and let j_n be an index for which $x_{j_n}^{(n)} = \max_{1\leq i\leq d} x_i^{(n)}$. Observe that for all n we have $x_{i_n}^{(n)} < c < d < x_{j_n}^{(n)}$. Since d is finite, here exist indices i and j and a subsequence n_k for which $i_{n_k} = i$ and $j_{n_k} = j$ for $k = 1, 2, 3, \ldots$, but then we have

   $$x_i^{(n_k)} - x_j^{(n_k)} > d - c, \text{ for } k = 1, 2, 3, \ldots$$

   which contradicts (i).

---

9. Let X_1 and X_2 be independent and identically distributed $Uniform(0,1)$ random variables. Let $Y = \min(X_1, X_2)$ and $Z = \max(X_1, X_2)$. Give the conditional density of Z given $Y = y$.

   Solution:

Note that $Y \leq Z$ with probability 1 and for $0 \leq y \leq z \leq 1$ we have

$$P[y \leq Y, Z \leq z] = P[y \leq X_1, X_2 \leq z] = \begin{cases} (z-y)^2 & \text{if } y \leq z, \\ 0 & \text{otherwise,} \end{cases}$$

Furthermore,

$$P[Z \leq z] = P[X_1, X_2 \leq z] = z^2,$$

for $0 \leq z \leq 1$, so we see that the joint cdf is

$$F_{Y,Z}(y,z) = P[Y \leq y, Z \leq z] = P[Z \leq z] - P[y \leq Y, Z \leq z] = \begin{cases} z^2 - (z-y)^2 & \text{if } 0 \leq y \leq z \leq 1, \\ z^2 & \text{otherwise,} \end{cases}$$

Then the joint density is given by

$$f_{Y,Z}(y,z) = \frac{\partial^2}{\partial y \partial z} F_{Y,Z}(y,z) = \begin{cases} 2 & \text{if } 0 \leq y \leq z \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

Thus (Y,Z) is uniform in the triangle $0 \leq y \leq z \leq 1$ and given $Y = y$, the distribution of Z is uniform in the interval $(y,1)$.

---

10. Prove that there exists an $n \times n$ real matrix A with $A^2 = -I$ if and only if n is even.

Solution:

$(\Rightarrow)$ If $A^2 = -I$, then

$$(-1)^n = \det(-I) = \det(A^2) = \det(A)^2$$

which is impossible if n is odd; so n must be even.

$(\Leftarrow)$ If n is even, set

$$A = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$$

where 0 is a block of zeros and I is an identity matrix, each of size $\frac{n}{2} \times \frac{n}{2}$. It is easy to see that $A^2 = -I$.

---

11. If X is a Poisson random variable with parameter λ, where $0 < \lambda < 1$, find $E[X!]$.

Solution:

$E[X!] = \sum_{i=0}^{\infty} i! \frac{e^{-\lambda}\lambda^i}{i!} = \sum_{i=0}^{\infty} e^{-\lambda}\lambda^i = e^{-\lambda}\sum_{i=0}^{\infty}\lambda^i = \frac{e^{-\lambda}}{1-\lambda}$.

---

12. Let V be the vector space of real polynomials with degree 2 or less. Consider the dot product on V
$$P \cdot Q = -\int_0^1 P(x)Q(x)\ln x dx.$$
   (a) Prove that this dot product is well defined and provides a positive definite linear form on V.

   (b) If the polynomial $Q = ax+b$ is orthogonal to $P \equiv 1$ what linear relation do the constants a and b satisfy?

Solution:

To prove that the dot product is well defined, it suffices to remark that the integral

$$\int_0^1 x^k \ln x dx$$

is well defined for $k = 0,\ldots,4$. If $k \geq 1$, it is the integral of a continuous function, and

$$\int_0^1 \ln x dx = [x \ln x - x]_0^1 = -1$$

is also well defined.

The associated form is positive definite since $(-\ln x) \geq 0$, implying that

$P \cdot P \geq 0$ and vanishes only if $P(x)^2 \ln x = 0$ on $(0,1)$ which implies $P = 0$ since P cannot have more than 2 roots without vanishing.

The inner product between $P \equiv 1$ and $Q = ax+b$ is given by

$$-\int_{x=0}^1 (ax+b)\ln(x)dx.$$

Integration by parts gives

$$\int_{x=0}^1 x \ln x dx = \frac{1}{2}x^2 \ln(x)|_{x=0}^1 - \frac{1}{2}\int_{x=0}^1 x dx = -1/4.$$

and using the fact that $\int_{x=0}^1 \ln(x)dx = -1$, we conclude that we get

$$P \cdot Q = a/4 + b,$$

so we must have $a/4 + b = 0$ if P and Q are orthogonal.

---

13. For positive integers n, m with $n < m$ define

$$f(n,m) = \frac{1}{n^2} + \frac{1}{(n+1)^2} + \cdots + \frac{1}{m^2}.$$

Prove that for every positive real number ε there must exist an integer T so that if $T \leq n < m$ then $f(n,m) < \varepsilon$.

Solution:

Define a sequence $z_1, z_2, \ldots$ by

$$z_n = \sum_{k=1}^{n} \frac{1}{k^2}.$$

It is well-known that $\{z_n\}$ converges and this can be readily confirmed by the integral test (because $\int_1^\infty x^{-2}\,dx$ converges). Therefore $\{z_n\}$ is a Cauchy sequence (a sequence converges iff it's a Cauchy sequence). Therefore, for every $\varepsilon > 0$ there is an integer T_0 so that for all $n, m > T_0$, $|z_n - z_m| < \varepsilon$. Wolog $n < m$ and then $|z_n - z_m| = f(n+1, m)$. Let $T = T_0 + 1$ and we have that $f(n,m) < \varepsilon$ for all $m > n \geq T$.

---

14. Let f be a nonnegative continuous function with $\int_{x=0}^{\infty} f(x)dx < +\infty$. Prove that

$$\lim_{n\to\infty} \frac{1}{n} \int_{x=0}^{n} x f(x)dx = 0.$$

Hint: Write the integral as a double-integral.

Solution:

We can write

$$\begin{aligned} I_n := \frac{1}{n}\int_{x=0}^{n} x f(x)dx &= \frac{1}{n}\int_{x=0}^{n}\int_{y=0}^{x} f(x)dydx \\ &= \frac{1}{n}\int_{y=0}^{n}\int_{x=y}^{n} f(x)dxdy \end{aligned}$$

Let $K = \int_{x=0}^{\infty} f(x)dx$. For any $\varepsilon > 0$ pick n such that $\int_{x=n}^{\infty} f(x)dx \leq \varepsilon/2$ and then choose $m \geq n$ so that $Kn/m < \varepsilon/2$. Then for $N > m$

$$\begin{aligned} I_N = \frac{1}{N}\int_{y=0}^{N}\int_{x=y}^{N} f(x)dxdy &= \frac{1}{N}\int_{y=0}^{n}\int_{x=y}^{N} f(x)dxdy + \frac{1}{N}\int_{y=n}^{N}\int_{x=y}^{N} f(x)dxdy \\ &\leq \frac{1}{N}\int_{y=0}^{n}\int_{x=0}^{\infty} f(x)dxdy + \frac{1}{N}\int_{y=n}^{N}\int_{x=y}^{\infty} f(x)dxdy \\ &\leq Kn/m + \frac{\varepsilon}{2}\frac{N-n}{N} \leq \varepsilon. \end{aligned}$$

---

15. The lifetime of a lightbulb has an exponential distribution with parameter 1. If n lightbulbs have independent lifetimes and all are turned on simultaneously, what is the expected value of the time at which the first bulb fails?

Solution:

Let T_i denote the lifetime of the i-th lightbulb, and let T denote the first failure time. Then so

$$P[T \geq t] = P[T_i \geq t, \text{ for } i = 1, \dots, n]$$

$$= P[T_i \geq t]^n = (e^{-t})^n = e^{-nt}.$$

So the cdf of T is $1 - e^{-nt}$ and the pdf is ne^{-nt} for $t > 0$. Now we calculate

$$E[T] = \int_{t=0}^{\infty} nte^{-nt} dt = 1/n.$$
