Department of Applied Mathematics
PROBABILITY AND STATISTICS PRELIMINARY EXAMINATION
August 2016

Notice: Do four of the following five problems. Place an X on the line opposite the number of the problem that you are **NOT** submitting for grading. Please do not write your name anywhere on this exam. You will be identified only by your student number, given below and on each page submitted for grading. Show all relevant work.

1.	____
2.	____
3.	____
4.	____
5.	____
Total	____

Student Number ___________________________

1. Suppose $X_1, X_2, \ldots$ are a sequence of iid U(0,1) random variables.

   (a) Find the distribution of $-\log X_1$ (log indicates the natural logarithm here).

   (b) Find the limiting distribution of

   $$\prod_{i=1}^{n} X_i^{-1/n}$$

   as $n \to \infty$.

   (c) Find the generic limiting distribution of a set of iid random variables $Y_1, Y_2, \ldots$.

   $$\prod_{i=1}^{n} Y_i^{-1/n}$$

   as $n \to \infty$. What assumptions are necessary for your answer?

2. Suppose (X, Y) are mean zero, variance one random variables with $\mathrm{Cov}(X, Y) = \rho$.

   (a) Find the constant a that minimizes $\mathbb{E}(Y - aX)^2$.

   (b) If (X_0, Y_0) has the same joint distribution as (X, Y), but is independent of (X, Y), find the mean and variance of

   $$aX + (Y_0 - aX_0).$$

   (c) Find the covariance between X and $aX + (Y_0 - aX_0)$.

   (d) If (X, Y) is additionally distributed as a bivariate normal, what can you say about the joint distribution of $(X, aX + (Y_0 - aX_0))$ as compared to (X, Y)?

(e) Suppose (X, Y) are jointly normal and you are given an observation $X = x$. How would you create a new observation y such that (x, y) was a realization from the distribution of (X, Y)?

3. Let $X_1, X_2, \ldots, X_n$ be a random sample, with $n \geq 3$, from the Poisson distribution with parameter θ. Consider estimating $\tau(\theta) = \theta^2 e^{-2\theta}$.

   (a) Find an unbiased estimator of $\tau(\theta)$. (*Hint: Consider only using X_1 and X_2.*)

   (b) Find the uniformly minimum variance unbiased estimator (UMVUE) of $\tau(\theta)$.

   (c) Find a non-trivial lower bound for the variance of any unbiased estimator of $\tau(\theta)$. ("Non-trvial" means don't use zero or a negative number!)

   (d) Find the MLE (maximum likelihood estimator) of $\tau(\theta)$.

   (e) You should have found that your MLE is different from your UMVUE. Is this alone enough to say that the variance of your MLE will not achieve your lower bound from part (d)?

4. Suppose ϕ is a random variable that defines a parameter of a random variable X in the exponential class of distributions. In particular, suppose

$$f(x|\phi) = c(\phi)h(x)e^{\phi t(x)}$$

   for some differentiable function $c(\cdot)$.

   (a) Take derivatives of $\int f(x|\phi)\mathrm{d}x = 1$ with respect to ϕ and show $\mathbb{E}(t(X)|\phi) = -c'(\phi)/c(\phi)$.

   (b) Suppose ϕ has a density of the form

$$f(\phi) \propto c(\phi)^n e^{n\theta\phi}.$$

   Similar to part (a), take the derivative of $f(\phi)$ and use the Fundamental Theorem of Calculus to find $\mathbb{E}t(X)$.

5. A strand of DNA for an organism is composed of a sequence of nucleotides corresponding to the nucleobases adenine (A), cytosine (C), guanine (G), and thymine (T). Consider the evolution of a single site which will always be in one of the states A, C, G, or T. Assume that, at each discrete time step, the state at the site evolves as a time-homogeneous Markov chain. In particular, assume that the nucleobase at time n changes to any particular different value with probability r $(0 \leq r < 1/3)$ or remains unchanged with probability $1 - 3r$.

   (a) What is the long-run probability of finding the site in state A?

(b) If the site starts in state A, what is the expected number of time steps it will take to return to state A?

(c) If the site starts in state A, what is the probability it will visit state T before first returning to state A?

(d) If the site starts in state A, what is the expected number of visits to state T before returning to state A?

(e) Consider the same DNA site for a separate, independently evolving organism. (Assume that the evolution of states at sites for this organism follows the same Markov chain as the first.) If the site for both organisms is in state A at time 0, find the probability that they are both in the same state at time n.

(f) Suppose that corresponding DNA segments of length N (N sites long) are being observed for both organisms. If both segments were the same at time zero, find the probability that they differ at exactly M sites at time n.