Bootcamp Challenge Problems

Duke StatSci Bootcamp, 2018

1 Day One: Calculus

1. Find the Taylor expansion of $f(x) = \tan x$ around 0 and use the result to show that $\lim_{x\to 0} \frac{\tan x}{x} = 1$.

2. Calculate $\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}$. Here $\lambda > 0$ is a constant and $x \in \mathbb{N}$. $(0! = 1.)$

3. Let $f(x) = x^{\alpha-1}(1-x)^{\beta-1}, x \in (0,1)$, where $\alpha, \beta > 0$ are constants. Find all the minimums and maximums (if there are any) of f.

4. Given that $\int_1^{\infty} \frac{1}{x}dx = \infty$ and $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$, use two methods to show that $\sum_{n=1}^{\infty} \frac{1}{n^p} = \infty$ for any $p \in (0,1)$.

5. Evaluate the following definite integrals:
   - $\int_0^1 \frac{2x}{1+x^4}dx$.
   - $\int_0^{\pi} e^x \sin x dx$.

6. The Gamma function $\Gamma(x)$ is defined for any real number $x > 0$ as $\Gamma(x) = \int_0^{\infty} t^{x-1}e^{-t}dt$. Show that $\Gamma(x+1) = x\Gamma(x)$.

7. Evaluate $\int_0^{\infty} e^{x^2}dx$.
   (Hint: start with $\int_0^{\infty}\int_0^{\infty} e^{x^2+y^2}dxdy$.)

8. Suppose X is a random variable that takes on only nonnegative integer values, with $E[X] = 1$, $E[X^2] = 2$, and $E[X^3] = 5$. (Here E[Y] denotes the expectation of the random variable Y.) Determine the smallest possible value of the probability of the event $X = 0$.

9. Find the smallest positive integer j such that for every polynomial $p(x)$ with integer coefficients and for every integer k, the integer

$$p^{(j)}(k) = \frac{d^j}{dx^j}p(x)|_{x=k}$$

   (the j-th derivative of $p(x)$ at k) is divisible by 2000.

10. Solve the previous problem, but for the general case where the number 2000 is replaced by n

2 Day Two: Matrix Algebra

1. Let A be a $m \times n$ matrix with rational entries. Suppose that there are at least $m+n$ distinct prime numbers among the absolute values of the entries of A. Show that the rank of A must be at least 2.

2. Fan and Kyle play the following game with an 2018×2018 matrix. Fan fills in one of the entries of the matrix with a real number, then Kyle, then Fan and so forth until the entire matrix is filled. At the end, the determinant of the matrix is taken. If it is nonzero, Fan wins; if it is zero, Kyle wins. Determine the optimal strategy of this game and who wins with perfect play.

3. Who wins in the game described in Problem 2 for a 3×3 matrix? A 5×5 matrix?

4. Let A be a $2n \times 2n$ matrix, with entries chosen at random. Each entry is chosen to be 0 or 1, each with probability 1/2. Find the expected value of $\det(A - A^\top)$ as a function of n, where $A^\top$ is the transpose of A.

3 Day Three: Probability and Distribution Theory

1. Kyle and Fan play a simple game of dice, as follows. Kyle keeps throwing the (fair) die until he obtains the sequence $\{1, 1\}$ in two successive throws. For Fan, the rules are similar, but she throws the die until she obtains the sequence $\{1, 2\}$ in two successive throws.

   (a) On average, will both have to throw the die the same number of times? If not, whose expected waiting time is shorter (no explicit calculations are required)?

   (b) Derive the actual expected waiting times for Kyle and Fan.

2. Kyle and Fan both want to cut out a rectangular piece of paper. Because they are both very nerdy, they determine the exact form of the rectangle by using realizations of a positive random variable, say U, as follows: Kyle is lazy and generates just a single realization of this random variable; he then cuts out a square that has length and width equal to this value. Fan likes diversity and generates two independent realizations of U. She then cuts out a rectangle with width equal to the first realization and length equal to the second realization.

   (a) Will the areas cut out by Kyle and Fan differ in expectation?

   (b) If they do, is Kyle's or Fan's rectangle expected to be larger?

3. Let M, N be independent Poisson random variables M and N have means α and β, respectively. (Recall that a Poisson random variable with mean λ has probability mass function $f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$). Define the random variable $Q = M^N$.

   (a) Find the expectation $E[Q]$.

   (b) What are the parameters α and β that maximize $E[Q]$, subject to the constraint that $\alpha + \beta = 10$?

4 Day Four: Inference

1. Kyle and Fan are given copies of the same bootcamp slides for independent proofreading. Fan finds 20 errors, and Kyle (who we learned yesterday is lazy) finds 15 errors, of which 10 were found by Fan as well. Estimate the number of errors remaining in the slides that have not been detected by either Kyle or Fan.

2. Let T be an exponential random variable with density $\lambda e^{-\lambda t}$. The parameter λ is called the rate of the density and $E[T] = 1/\lambda$. Exponential random variables have many interesting properties, two of which are listed below.

   (i) The minimum of independent exponential random variables, which may have different rates, is again exponentially distributed, with a rate equal to the sum of the individual rates. For example, the minimum of two independent exponential random variables with rates λ_1 and λ_2 is again an exponential random variable, with rate $\lambda_1 + \lambda_2$.
   (ii) If an exponential random variable T is larger than some other positive independent random variable Q, then the difference $T - Q$ is again exponentially distributed with rate λ. This is sometimes called the "memoryless' property of the exponential distribution.

   We draw a sample of size $n = 2k - 1$ (k a positive integer) i.i.d. realizations of T, and estimate the median of T by the kth order statistic of this sample, i.e., the midmost element of the sample that is both larger and smaller than $k - 1$ other realizations. Call this median estimate M.

   (a) Derive the density of M.
   (b) Find an expression for the expectation of M, and show that for small k it is severely biased.