CS 3190: Foundations of Data Analysis **Due: Tuesday 9.28 at 11:50pm**

Homework 2: Convergence and Linear Algebra

Instructions: Your answers are due at 11:50pm submitted on canvas. You must turn in a pdf through canvas. I recommend using latex (`http://www.cs.utah.edu/~jeffp/teaching/latex/`, see also `http://overleaf.com`) for producing the assignment answers. If the answers are too hard to read you will lose points, entire questions may be given a 0 (e.g. **sloppy pictures with your phone's camera are not ok, but very careful ones are**)

Please make sure your name appears at the top of the page.

You may discuss the concepts with your classmates, but write up the answers entirely on your own. **Be sure to show all the work involved in deriving your answers! If you just give a final answer without explanation, you may not receive credit for that question.**

---

1. [**30 points**] Consider a random variable X with expected values $\mathbf{E}[X] = 10$ and variance $\mathbf{Var}[X] = 2$. We would like to upper bound the probability $\mathbf{Pr}[X > 12]$.
   (a) Which bound can and cannot be used with what we know about X (Markov, Chebyshev, or Chernoff-Hoeffding), and why?
   (b) Using that bound, calculate an upper bound for $\mathbf{Pr}[X > 12]$.
   (c) Describe a probability distribution for X where the other two bounds are definitely not applicable.

2. [**30 points**] Consider a pdf f so that a random variable $X \sim f$ has expected value $\mathbf{E}(X) = 2$ and variance $\mathbf{Var}(X) = 4$. Now consider $n = 10$ iid random variables $X_1, X_2, \ldots, X_{10}$ drawn from f. Let $\bar{X} = \frac{1}{10}\sum_{i=1}^{10} X_i$
   (a) What is $\mathbf{E}(\bar{X})$?
   (b) What is $\mathbf{Var}(\bar{X})$?
   Assume we know X is never smaller than 0 and never larger than 7
   (c) Use Markov inequality to upper bound $\mathbf{Pr}(\bar{X} > 3)$
   (d) Use Chebyshev inequality to upper bound $\mathbf{Pr}(\bar{X} > 3)$
   (e) Use Chernoff-Hoeffding inequality to upper bound $\mathbf{Pr}(\bar{X} > 3)$
   (f) Now suppose $n = 100$. Calculate the 3 bounds again. For this part, also make sure to report the *name* of the inequality that gives the tightest bound, the second tightest bound, and the third tightest bound respectively.

3. [**15 points**] Consider the following 2 vectors in $\mathbb{R}^4$:

$$\begin{aligned} p &= (2, -4, 8, \mathtt{x}) \\ q &= (4, -8, 16, -10) \end{aligned}$$

   Report the following:

(a) Choose the value `x` so that p and q are linearly dependent

(b) Choose the value `x` so that p and q are orthogonal

(c) Choose a single value of `x` so that p and q are neither linearly dependent nor orthogonal

(d) Calculate $\|q\|_1$

(e) Calculate $\|q\|_2^2$

4. **[25 points]** Consider the following 2 matrices:

$$A = \begin{bmatrix} 2 & -1 & 4 \\ 0 & -1 & 0 \\ 3 & -2 & 6 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Report the following (e.g using Python):

(a) $A^T B$

(b) AB

(c) BA

(d) $B + A$

(e) B^T

(f) Which matrices are invertable? For any that are invertable, report the result.