Statistics 5114: HW 0

This assignment will help you get over some of the calculus and probability hurdles that often present themselves in class. We won't require formal proofs in class, but being able to prove things formally is a useful skill. Even drawing pictures can be acceptable if the ideas are right, and the concepts are reasonably well explained. During class, I'll make clear what needs to be proved/etc. Enjoy!!! **Select 5 problems of your choosing and write up solutions (neatly; LaTeXis preferred). You'll turn these in (on paper) on the due date assigned**.

1. Show that $lim_{n \to \infty}(1 + \frac{1}{n})^n \to e$.

2. Show that $lim_{n \to \infty}(1 + \frac{f(x)}{n})^n \to e^{f(x)}$.

   Note that these result can be shown several different ways, and in some classes is taken as "definition". Just review the result and know it:) A math major might disagree with a proof being correct/incorrect, but we won't be that rigid in class.

3. Review Jensen's inequality which shows a relationship between $E[g(x)]$ and $g(E[x])$ (either $\geq$ or $\leq$), where $g(\cdot)$ is either a convex or concave function.

4. Prove that the inequality switches to an equality in Jensen's inequality if and only if $g(\cdot)$ is a linear function. Note that you don't need to use Jensen's to prove that, since $E[x]$ is a "linear operator", but this will help you to think through Jensen's.

5. Let x_1 and x_2 have a bivariate Normal distribution, which can be expressed by:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right),$$

   where $\Sigma_{21} = \Sigma_{12}^T$. Note that I'm being really sloppy about the dimensionality of $x_1, x_2, \mu_1, \mu_2, \Sigma_{11}, \Sigma_{12}, \Sigma_{21}$, and Σ_{22}. I'd like you to label the dimensions and make sure everything is "conformable" (look that word up if you don't know it:)).

   Find the conditional distribution for $x_1|x_2$, which can be verbally expressed as x_1 given x_2. Note that when we say the words "given" x_2, x_2 is fixed/constant.

6. Prove that the variance for $x_1|x_2$ is smaller or equal to the marginal variance of x_1.

7. Relate the mean/expected value of $x_1|x_2$ to what you have learned in your Regression/Anova class.

8. This is a new problem, and the probability distribution does not matter. Prove that $f(x_1|x_2) = f(x_1)$ is an equivalent to saying $f(x_1, x_2) = f(x_1)f(x_2)$ (i.e x_1 and x_2 are independent).

9. Review Bayes' Theorem.

10. Review likelihood functions. Here's a result: if $x_i \sim N(\mu, \sigma^2)$ $(x_i' sare independent$ and let σ^2 is a fixed known constant,; Note that if I don't say they are independent in my upcoming class, and I don't explain any dependence "mechanism", you can assume everything is independent:)), the likelihood function for μ is $\prod_{i=1}^{n} e^{-\frac{1}{2}(x_i-\mu)^2/\sigma^2}$, which a function of μ (this is not a function of the $x_i's$ since for likelihood functions "condition" on the data. That is, the $x_i's$ are considered fixed/observed).

    Why did I drop the leading term $\frac{1}{\sqrt{2\pi\sigma^2}}$?

11. Show that the above likelihood function is equivalent to $Ce^{-\frac{1}{2}(\mu-\bar{x})^2/(\sigma^2/N)}$, where C can be any positive constant.

12. Review Taylor's Series/Power Series, which states something like:

$$f(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + a_3(x-a)^3 + ... + a_n(x-a)^n,$$

    where the coefficients are:

$$a_n = \frac{f^{(n)}(a)}{n!}.$$

13. What does Taylor's Theorem say about the remainder term? What is the "radius of convergence"?

    Let's spend a little effort to get ahead on some problems:) This is optional.

14. This is probably a new result for you, but it's worth studying:) Use Google/Stack Exchange/Quora to help you look up the terms and help you through the result. This is not in Casella and Berger.

    Prove that if $\theta \sim p(\theta)$ and $Y_1, \ldots, Y_n$ are conditionally i.i.d. given θ, then marginally (unconditionally on θ), $Y_1, \ldots, Y_n$ are exchangeable. I like this result a lot:)

15. An assignment that usually gives students a bit of trouble is my EM algorithm project which can be found at (https://leman.stat.vt.edu/VTCourses/5114EM.pdf). Have a look at Casella and Berger, and see if you can understand the algorithm. The explanation in the book is a little difficult to follow. Have a look at my video series on this, which can be found starting from the link (https://www.youtube.com/watch?v=zrZjZ7Sp6-Q). I have lots of videos on that YouTube page, so if you're really looking to get ahead, you can scroll through my previous lectures.

16. Have a look at last years HW assignments, which can all be found at: https://leman.stat.vt.edu/VTCourses/5114Schedule.html