Math 266, Probability & Statistics, Spring 2013, Prof. Janeba
Project #2 - the pdf for $X_1 + X_2 + X_3$ where X_i is uniformly distributed on $[0, 1]$

Your assignment: Use the outline and hints below to write a self-contained paper pertinent to the title of this project and the paragraph below that cites the Central Limit Theorem.

Self-contained means that the point of the paper is completely clear and understandable to someone who has not read this prompt (at all).

Since this prompt is slightly cryptic about the point of it all, part of your assignment is to discern (and explain) that as well. What does this lengthy investigation demonstrate? Do not underestimate this part of the project. Of course, you can't start this part until you've finished all the computations...

For absolute top honors, your paper would need to include some sort of counterpoint to the conclusion - just because something happens here for a small value of n, does that mean it always happens for small n?

We know from the Central Limit Theorem that when we add n independent instances of a random variable X, i.e. add many identically-distributed random variables, that when n is sufficiently large, the distribution for the sum is approximately normal. This project will explore just how closely the distribution can come to being normal, even when n is relatively small.

Our primary method will be to use joint probability density functions for two variables at a time. Read chapter 7.2 of our text, paying particular attention to example 7.3. Really, now. Go read it.

Are you back? Looking more closely at example 7.3 we have $f(x_1, x_2) = 6e^{-3x_1-2x_2}$ when x_1 and x_2 are positive; for such values, $f(x_1, x_2) = 6e^{-3x_1-2x_2} = 6e^{-3x_1}e^{-2x_2} = \left(3e^{-3x_1}\right)\left(2e^{-2x_2}\right)$ which just happens to equal $f_1(x_1) \cdot f_2(x_2)$, the product of the two marginal density functions. Looking at Definition 3.14 on p. 111, we see that this means X_1 and X_2 are independent.

So let's get down to *this* project. Define X_1, X_2, and X_3 as independent uniformly distributed random variables on $[0, 1]$. (See section 6.2).

Leaving X_3 aside for the moment, use the method of example 7.3 to find the cumulative distribution function for $Y = X_1 + X_2$ (we are defining Y here). Note the joint pdf for X_1, X_2 is just the product of the pdfs of X_1 and X_2. Notice this problem will differ from example 7.3 in that the limits of integration will be quite different; while the example problem has a nonzero joint pdf for $x_1 > 0, x_2 > 0$, this problem is more constrained - we need $0 < x_1 < 1$ and $0 < x_2 < 1$ to get a nonzero joint pdf. This will make setting up the double integral a bit trickier.

Here are two final hints for finding the cdf for Y : First, if you set up integrals, you will want to take the case $y < 1$ separate from the case $y > 1$ in finding $F(y)$. Secondly, symbolic integration can be avoided altogether for this part (your option) if you think of the double integral as the volume of some solid.

Next, differentiate the cdf for Y to get a pdf for Y. Check both the cdf and pdf for the properties that all cdfs and pdfs must have, and whether the specific answer is plausible to you. Don't forget to graph the pdf to see if it tells us something.

Repeat (lather & rinse optional). That is, find a joint pdf for Y and X_3 (remember him, X_3?). Y and X_3 are independent, and we now have their separate pdfs, so this should be easy. Use this joint pdf to find the cumulative distribution function for $W = Y + X_3$,and then differentiate it. As you do, modify the hints and methods above. Then explain what the result has to do with the title of the project.

(Over for **The Rules**).

The Rules: Do your work in groups of two or three, turn in one paper, typed, with all group member's names on it. Note that you must explain your answer in great (great!!) detail, so that you are not only right, but your paper will persuade a fellow student whose answer is not right. This means there will be far, far more words in your paper than equations or math symbols, though there will be plenty of those too. Because you are Willamette students, the expectation is that you will use grammatically and orthographically correct, complete sentences, seamlessly interwoven with your math symbols. See the syllabus for the (perhaps unexpected) rules about academic honesty for written group assignments. Due by Tuesday, April 16, at 4:00 p.m. at my office.

This mini-project is worth 50 class points. It is also an opportunity for those who feel they don't or haven't tested well to raise their grades.

Group evaluation: Your group turns in one report and everyone in the group gets the same grade. To provide some accountability, however, every group member is required to submit to an informal group evaluation. Handwritten, typed, or email are OK. The evaluation should include (just!) a few sentences about how well the group worked together. The evaluation must include your estimate of the percentage of the project work that you feel each group member did (e.g. 33%/33%/33% or 40%/35%/35%, etc.). These percentages must add to 100%. These evaluations will be held in confidence, and used at the end of the semester, if necessary, to adjust final grades.