PS#14 - The normal distribution!

The well-known "68-95-99.7 rule" states that approximately 68% of values of a normally-distributed random variable lie within one standard deviation of the mean. In this extended problem, we will use Taylor series to calculate this percentage directly.

Here is the equation for the "normal probability density function" with mean $\mu = 0$ and standard deviation $\sigma = 1$:

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right).$$

You may know this function better by another name: the bell curve!

Throughout this problem set, when you are asked to write down a Taylor series for some function, you may either use Σ notation or write out the first few terms of the series.

1. Draw a graph of $p(x)$. You can use Desmos but in a minute I'm going to ask you to shade something in so you may want something you can draw on.

2. We want to find the proportion of values that lie within one standard deviation ($\sigma = 1$) of the mean ($\mu = 0$); in other words, we want to find the proportion of values that lie less than 1 distance away from 0 in either direction. Shade in this region on your graph.

3. We can compute this value by finding $\int_a^b p(x)\,dx$. What should our bounds of integration a and b be? Why?

4. If we find an antiderivative, $P(x)$, then our answer will be $P(b) - P(a)$. But unfortunately, we can't immediately write down an antiderivative $P(x)$. Why not?

5. Write down the Taylor series for $f(u) = e^u$.

6. Use the Taylor series for e^u to write down the Taylor series for $p(x)$. (Don't forget the $\frac{1}{\sqrt{2\pi}}$!)

7. Use the Taylor series for $p(x)$ to write down a Taylor series for its antiderivative $P(x)$.

8. Use the first four nonzero terms of the Taylor series for $P(x)$ to approximate $P(b)$.
(Hint: This is probably *not* the fourth-degree polynomial.)

9. Use the first four nonzero terms of the Taylor series for $P(x)$ to approximate $P(a)$.
(Notice anything interesting about $P(a)$ in relation to $P(b)$?)

10. Find $P(b) - P(a)$. (Your answer should be close to 68%.)

11. How many terms would be necessary to ensure that your approximation to $P(b) - P(a)$ is accurate to four decimal places? **Notes:**
    - By "accurate to four decimal places," we mean that if we rounded both our approximation and the actual value to the fourth decimal place, they'd round to the same thing. How far off can we be?
    - You can't compare your computed value directly to the actual value, because the actual value is unknown!
    - This is an alternating series, so we can use the alternating series error bound.
    - We're actually using two values, $P(b)$ and $P(a)$, so we can accrue some error from both ends. You may have noticed something nice about how $P(b)$ and $P(a)$ are related that will be helpful.

12. Conclude by telling me why Taylor series are really awesome for dealing with yucky integrals.

13. **(Bonus!)** Similarly, check the "95" and "99.7" parts of the rule, which talk about the proportion of values within 2 and 3 standard deviations of the mean. (What will you have to change in your previous work?)