INSTRUCTIONS: Solve any 4 of the following 5 problems, clearly labeling each solution. Write your solutions clearly and use English words and sentences to explain your work where appropriate. Even a correct solution may receive little or no credit if a method of solution is not shown/described ("I used R" may be sufficient). You may use a computer, notes, your book, and any other hon-human resources. You may also ask Dr. Axon for hints, but you may not discuss these questions with anyone else. Solutions are due by 5:00 PM on Friday, May 6.

1. The authors of the article "Boredom in Young adults–Gender and Cultural Comparisons" administered the Boredom Proneness Scale to 97 male and 148 female U.S. college students. A summary of the results is shown below (a higher score means increased tendency towards boredom). Let μ_1 be the mean boredom proneness of male college students and μ_2 be the mean boredom proneness of female college students. Test the hypothesis $H_0: \ \mu_1 - \mu_2 = 0$ against the alternative hypothesis $H_1: \ \mu_1 - \mu_0 \neq 0$.

Gender	Sample Size	Sample Mean	Sample SD
Male	97	10.40	4.83
Female	148	9.26	4.68

a) Give the P-value for your hypothesis test.

b) State a clear conclusion (reject or fail to reject the null hypothesis) for testing at a significance level of $\alpha = 0.05$.

2. In a study of 5 triceratops skulls the brow horns had lengths (in cm) of 49, 53, 60, 41, and 82. Calculate a 90% confidence interval for the mean length of a triceratops brow horn.

3. A random sample of 34 bags of fun-sized M&Ms was analyzed (and then eaten). The numbers of green and yellow M&Ms in each bag was recorded: the results are at `http://web02.gonzaga.edu/faculty/axon/321/yellow-green.txt`. Fit a least squares line relating the numbers of green and yellow M&Ms and give the equation of the line.

4. Let $\overline{X}$ be the mean of a random sample of size n from an population uniformly (continuously) distributed on the interval $(\alpha, 1)$. Find constants a and b such that $a\overline{X} + b$ is an unbiased estimator for α.

5. An epidemiologist is trying to discover the cause of a certain kind of cancer. He studies a group of 10,000 people for five years, measuring 48 different "factors" involving eating habits, drinking habits, exercise, and so on. His object is to determine for each factor if there is a difference between the mean value of the factor among those who develop cancer and among those who do not: a difference in means will suggest a connnection between the factor and the cancer. He will test the null hypothesis $H_0 : \mu_1 - \mu_2 = 0$ against $H_1 : \mu_1 - \mu_2 \neq 0$ for each of the 48 factors (48 hypothesis tests in total). In an effort to be cautiously conservative, he uses a significance level of $\alpha = 0.01$ in all his tests. What is the probability that at least one of the factors will be found to be associated with the cancer, even if none of them is actually connected (that is, what is the probability of making at least one type I error)?