Problem Set 1
ECON 30331

(Problems 1, 2, 3, 4, 8, 10, and 11 are due at the start of class, Tuesday, Sept. 1st)
Dan Hungerman
Spring 2009

1. The fraction of impurity (x) in a metal generated by a production process is given by the PDF f(x) = 2x for $0 \le x \le 1$.
   a. Show that f(x) is a proper pdf, that is, it integrates to 1 over the range [0,1]?
   b. What is E[x]?
   c. What is Prob[X<0.25]?

2. Suppose the time (in years) until a car fails (x) can be described by the exponential distribution, $f(x) = \lambda e^{-\lambda x}$, for $\lambda>0$ and x>0. Suppose $\lambda=0.25$ and the car warranty is for 2 years. What is the probability the car will fail *after* the warranty has expired?

3. Suppose that the random variable (z) is distributed as a standard normal distribution. Using a standard normal table, calculate the following probabilities:
   d. Prob($Z \le -1.65$)
   e. Prob($Z > 1.25$)
   f. Prob($-1.02 < z \le 0.53$)

4. Students who graduate in the top 10 percent of their high school class from an accredited Texas school are guaranteed admission to the University of Texas at Austin. Class rank is typically determined by cumulative grade point average (GPA). Assume that the cumulative GPA for the class of 2007 at Tom Landry High School from Arlen, Texas is normally distributed with a mean of 2.7 and a standard deviation of 0.5. What GPA must a student from the Tom Landry class of 2003 receive in order to be admitted to the University of Texas?

5. Listed below are 10 values from two series. Without calculating values, what is the correlation coefficient for x and y (hint: graph the points).

X	10	12	14	16	18	20	22	24	26	28
Y	20	19	18	17	16	15	14	13	12	11

6. Temperatures in Fahrenheit (F) can be converted to Celsius (C) using the simple linear transformation C = -17.78 + 0.556F. Show that the correlation coefficient between the average daily highs in temperature for a given city measured in Fahrenheit and Celsius is $\rho_{fc} = 1$.

7. Applicants to San Diego State University are conditionally accepted for admission if they have an "Eligibility Index" in excess of 4000. The Index (I) is a linear combination of the applicant's high school GPA and total SAT score where I = 800GPA + SAT. Suppose the applicant pool to the school has the following characteristics: The expected values of GPA and SAT are 3.6 and 1100 respectively, the standard deviations of GPA and SAT are 1 and 100, respectively, and the **correlation coefficient** between GPA and SAT is 0.5. What is the expected value and variance of the "Eligibility Index" for the population of applicants to this school?

   Bonus question: What fraction of applicants will be admitted?

8. One of the most studied areas of social science research is the impact of socioeconomic status on health and mortality. This question evaluates some of the evidence in this area. Listed below are three facts about mortality and poverty rates for adults males aged 24-64 in 1990.

- 14 percent of these people are in poverty, Pr(Poverty) = 0.14
- 3 percent of these people will die over the next five years, or Pr(Die on 5 years) = 0.03.
- 1 percent of these people will both die sometime over the next five years and are currently in poverty, or, Pr(Die in 5 Years $\cap$ In Poverty) = 0.01.

A. What is the Pr(Die in 5 years | In Poverty)?
B. What is the Pr(Die in 5 years | NOT in Poverty)?
C. Using the results above, determine whether ‘Probability of dying in the next five years’ is independent of ‘poverty status’.

Hint: draw the (2 x 2) box that is on page 9 of the “What I Should Have Learned in ECON 30330” Handout.

9. Researchers took the body temperatures of 25 healthy adults and obtained an average temperature (x) of 98.1 degrees Fahrenheit with a standard deviation (s) of 1.0 degree. For this sample, what is the 95% confidence interval for expected body temperature (μ) among healthy adults? With these results, test the null hypothesis that expected body temperature is equal to 98.6 degrees, H_o: μ=98.6.

10. A popular weight loss program these days is the ‘Atkins Diet’ that stresses a high protein/low carbohydrate diet. The Atkins diet is in stark contrast to many diets that stress low fats and /high carbohydrates. To examine which diet produces greater weight loss, a group of researchers randomly assigned 40 overweight patients to either an Atkins type diet (n_A=20) or a Low Fat/High Carb diet (n_L=20). Participants were given detailed guidelines about the diets and a phone number for a ‘diet hot line’ to call with any questions about the program they were assigned. The participants were weighed at the beginning of the diet and again in 6 months. After 6 months, dieters in the Atkins plan lost an average of 16 pounds whereas the participants in the Low Fat/High Carb diet lost only an average of 7 pounds. A summary of results from the experiment are listed below.

	Atkins Diet	Low Fat/High Carb Diet
n (sample sizes)	20	20
x (average change in weight, in pounds)	-16	-7
S	12	8

A. Construct a **95% confidence interval** for d =u_A - u_L-- the difference in weight loss generated by the Atkins and Low Fat diets. What are the appropriate degrees of freedom and t-value you should use in this context?
B. Using the confidence interval from a), test the null hypothesis that H_o: d=0 -- there is no difference in the weight loss generated by the Atkins versus the Low Fat diets. Can you reject or not reject the null hypothesis? Explain your answer.
C. Using a **t-test** and a **99%** confidence level, test the null hypothesis that H_o: d=0 -- there is no difference in the weight loss generated by the Atkins versus the Low Fat diets. What are the appropriate degrees of freedom of the t-statistic you use in this context? Can you reject or not reject the null hypothesis? Explain your answer.

11. Suppose Y_1 and Y_2 are independent, $var(Y_1) = Var(Y_2) = \sigma^2_y$ and $Z_1 = Y_1 + Y_2$. What is Var(Z)?

12. Problem 11 continued -- generalize the results of problem 11. Suppose $Y_1, Y_2 .... Y_n$ are independent, $var(Y_1) = Var(Y_2) = ...Var(Y_n) = \sigma^2_y$ and $Z_2 = Y_1 + Y_2 + Y_3 + ...Y_n$. What is $Var(Z_2)$?