Extra Practice Problems

True/False

For each of the following questions indicate true or false, then **explain** your answer. You may use examples to illustrate your answer.

(I) For two events A and B (where $Pr\{B\} > 0$ and $Pr\{A\} > 0$), $Pr\{A|B\}$ must be greater than or equal to $Pr\{A\}$.

(II) If a 95% confidence interval for β_1 is: (5.6, 9.4), it is appropriate to interpret this interval as "There is a 95% chance that the true slope β_1 is in the interval 5.6 to 9.4".
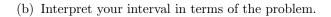
(III) One of the assumptions for a binomial random variable is that the probability of success remains constant (the same value) for all n trials.

(IV) If we wanted to minimize the probability of a Type II error, we should pick the largest value of α possible.

(V) The larger $(o_i - e_i)^2$ is for a χ^2 Goodness of fit test, the larger the p-value for the two-sided test will be. You may assume the number of categories is 2.

(VI) If the sample correlation is zero, that means the estimated slope for the regression line will also be zero.

1. A study to determine the effectiveness of a drug for arthritis was conducted, with 200 arthritic patients randomly put in a placebo group or a drug group. They were later asked if their arthris improved:

	Improved	No Improvement
Drug Group	117	83
Placebo Group	72	129

(a) Find a 95% confidence interval for the difference in the probability of improvement for the drug vs placebo group.

(b) Interpret your interval in terms of the problem.

(c) Does your interval from (a) suggest that improvement is independent of group? Explain.

(d) List all the necessary assumptions for a confidence interval for a difference in two proportions.

2. The following data summarizes the incidence of Coronary Heart Disease (CHD) for people who smoked regularly and for people who did not smoke regularly:

	CHD	No CHD
Smoked	84	87
Did not Smoke	296	491

For the hypothesis $H_0 : Pr\{CHD|Smoke\} = Pr\{CHD|Smoke^C\}$, the p-value was: 0.00528.

(a) Find the estimated probability of having CHD if you smoke, and the estimated probability of having CHD if you do not smoke.

(b) If the goal is to test if smoking status is independent of CHD, state the alternative hypothesis and find the p-value.

(c) What value of α (out of the three 0.01, 0.05, and 0.10) should we pick if we want the probability of a type I error to be as small as possible? Explain your answer.

(d) Estimate the probability that someone has CHD.

3. 60 guinea pigs were measured, and the amount of vitamin C (in mg) and the length of their adult teeth (in mm) was measured. We believe that vitamin C may increase the length of their teeth. Summary statistics are:

	Vitamin C (mg)	Tooth Length (mm)
sample mean	1.17	18.81
sample std. dev	0.63	7.65

In addition, $r = 0.803$, and $SS(Resid) = 1227.905$.

(a) Find the estimated regression line.

(b) Find the 95% confidence interval for the slope, and interpret it in terms of the problem.

(c) If a guinea pig that received 2 mg of vitamin C has a tooth length of 29.4 mm, find the error based on your regression line. Did the line over estimate or under estimate the actual value?

(d) List two of the assumptions for linear regression.

(e) Find and interpret s_e.

(f) Find and interpret r^2.

4. A biology professor believes the proportion of males taking their class is 60%, and the proportion of females taking their class is 40%. For a class of size 489, she finds that 223 are female, and 266 are male.

(a) State the appropriate null and alternative for the professors belief.

(b) Calculate the test-statistic for the hypothesis in (a).

(c) Calculate the p-value for the test in (a).

(d) State your conclusion in terms of the problem, using $\alpha = 0.01$.

(e) Were there more or less females taking the class than the professor expected? Explain.