Prof. David Draper
Department of Applied Mathematics and Statistics
University of California, Santa Cruz

AMS 131: Quiz 3

Name: _____

In the spring of 1993 I taught an introductory statistics class at UCLA. One of the things I did to generate data for analysis in the class was to conduct a (voluntary) survey of the students at the beginning of the quarter: I asked some demographic questions, including gender, and some political questions, including "Are you in favor of the legalization of marijuana?" Let's agree to code the gender variable as Female (F) or Male (M), and the marijuana legalization preference (MLP) variable as Yes (Y) or No (N). A total of 106 students responded to the survey; the results are summarized in the table below.

	MLP		
Gender	Y	N	Total
F	29	20	49
M	52	5	57
Total	81	25	106

In other words, 29 Female students said Yes (upper left cell), and there were a total of 25 people who said No (second column total).

In parts (a), (b) and (c), if a student is chosen at random from these 106 survey participants,

(a) What's the probability $P(Y)$ that the chosen person responded Yes to the MLP question? Explain briefly (for example, the right denominator for this probability is m because ..., and the right numerator is n because ...).

(b) Given that the chosen person is Female, what's the conditional probability $P(Y \mid F)$ that she responded Yes? Explain briefly.

(c) Given that the chosen person is Male, what's the conditional probability $P(Y \mid M)$ that he responded Yes? Explain briefly.

(over)

(d) Briefly explain why this demonstrates that gender and marijuana legalization preference are (probabilistically) *dependent* in this data set, and briefly describe the nature of the dependence. (*Hint:* What would have been true if these two variables had been *independent*? Think like a Bayesian.)

(e) Would you describe the degree of dependence in (d) as weak, moderate or strong? Use your results in parts (a), (b) and (c) to justify your answer.