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

AMS 131: Quiz 3

Name:
Jeremy Lafond

1. 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 ...).

$$\text{P(Y)} = \frac{\text{number of outcomes favorable to Y}}{\text{total number of possible outcomes}} = \frac{81}{106}$$

The right numerator is 81 because this number represents the total number of YES answers. The right denominator is 106 and it represents the total number of students that responded to the survey.

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

$$P(Y \mid F) = \frac{29}{49}$$

This probability is a result of the fact that out of 49 females, 29 of them answered yes on the survey.

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

$P(Y \mid M) = \frac{52}{57}$

This probability is a result of the fact that out of 57 males, 52 of them answered yes on the survey.

(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.)

Two events, A and B, are independent if $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$

$P(Y) = \frac{81}{106} \approx 76\%$ answered YES on MLP.

$P(M) = \frac{57}{106} \approx 54\%$ of those surveyed were MALE.

$P(Y \mid M) = \frac{52}{57} \approx 91\%$ of MALES answered YES on MLP.

$P(M \mid Y) = \frac{52}{81} \approx 64\%$ of YES answers on MLP were MALE.

$P(Y \mid F) = \frac{29}{49} \approx 59\%$ of FEMALES answered YES on MLP.

$P(Y \mid M) \neq P(Y)$ and $P(M \mid Y) \neq P(M)$

This shows that gender and MLP are not independent because knowing a random person is a male increases the probability that their MLP is YES. Therefore, there is a degree of dependence.

(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.

$P(Y) = \frac{81}{106} \approx 76\%$ answered YES on MLP.

$P(Y \mid M) = \frac{52}{57} \approx 91\%$ of MALES answered YES on MLP.

$P(Y \mid F) = \frac{29}{49} \approx 59\%$ of FEMALES answered YES on MLP.

$P(M \mid Y) = \frac{52}{81} \approx 64\%$ of YES answers on MLP were MALE.

$P(F \mid Y) = \frac{29}{81} \approx 59\%$ of FEMALES answered YES on MLP.

$\frac{\texttt{52 YES MALES - 29 YES FEMALES}}{\mathbf{81}} \approx 0.28$

Means that of those surveyed, you are approximately 28% more likely to be MALE if you answered YES on MLP

The degree of dependence in this relationship would then fall just below moderate and technically be weak if we considered 1-33% weak, 33-66% moderate, and 66-100% strong.