STA 113 Spring 2005
I. H. Dinwoodie
Testing for Hardy-Weinberg Equilibrium

A table of genotypes at a single location (locus) with four alleles A_1, A_2, A_3, A_4 is given by

	A_1	A_2	A_3	A_4
A_1	0			
A_2	3	1		
A_3	5	18	1	
A_4	3	7	5	2

This table classifies 45 individuals by the observed pair of alleles. The data comes from a paper of Louis and Dempster (1987) in the journal *Biometrics*. We want to know if the alleles are combining independently with one another.

1. Derive the probability of the given table in terms of p_1, p_2, p_3, p_4, the individual allele probabilities (or relative frequencies in the population), assuming they combine independently to produce genotypes (unordered pairs). For this you need the multinomial distribution on 10 categories corresponding to cells in the table. The probability that an individual turns out to be A_1/A_1 is p_1^2, but the probability that an individual turns out to be A_1/A_2 is $2p_1p_2$.

2. Find the maximum likelihood estimates for the probabilities p_1, p_2, p_3, p_4.

3. Use the χ^2 test (in our book on p. 644) to see if the table deviates too much from the closest table under Hardy-Weinberg assumptions. Be sure to get the right degrees-of-freedom (there are three free parameters in the model), and report a p-value for the test:

   H_0 : The population is in Hardy-Weinberg equilibrium.
   H_a : The population is not in Hardy-Weinberg equilibrium.