Stat 135, Fall '00, FINAL EXAM

1. In a 1974 survey, respondents were asked whether they thought it was all right for a woman to have a job instead of taking care of the house and children while her husband worked. The table shows the response by ethnic origin of the respondent. Is there a relationship between response and ethnic group?

	Yes	No
English	53	32
Irish	43	30
French	42	23
Italian	78	47
Portuguese	29	7

2. Suppose $X_1, \ldots, X_n$ are i.i.d. with density

$$f(x) = b(1-x)^{b-1} \qquad 0 < x < 1.$$

Find the method of moments estimate of b.

3. In a large introductory statistics class, we have the following statistics on midterm and final scores.

	AVG	SD
Midterm	65	10
Final	65	10

The correlation is positive and the distributions are approximately normal.

Consider the following statement:

> For those students who score 55 on the midterm, we expect half to score over 55 on the final.

Is this statement true or false? Or do you need more information to determine whether it is true or false. Explain (use calculation to support your explanation).

4. For $x_1, \ldots, x_n$ i.i.d. with density,

$$f(x) = \lambda e^{-\lambda x} \qquad x > 0$$

Find the asymptotic variance of the Maximum Likelihood estimate for λ. (No need to find the MLE).

5. For a large sample of Scottish soldiers, their chest size in inches has a histogram that looks normal. The 20th percentile is 38.3 inches, and the 60th percentile is at 40.5 inches.

Provide an estimate for the average chest size and the SD of chest size.

6. In a tape drive experiment, one tape is measured 6 times on each of 10 tape drives. The jth measurement on the ith tape is

$$Y_{i,j} = \mu + A_i + E_{i,j}$$

where $\mathbf{E}(A_i) = 0$, $\mathrm{Var}(A_i) = \sigma_A^2$,
$\mathbf{E}(E_{i,j}) = 0$, $\mathrm{Var}(E_{i,j}) = \sigma^2$, and all A_i and $E_{i,j}$ are independent.

a) Compute the variance of $Y_{i,j}$.
b) The one way analysis of variance is fit to the data. Find the expected value of the MSE.

7. Twenty measurement $y_1, \ldots, y_{20}$ are taken. The first 10 are replicates measured at x_1, and the next 10 are replicate measurements at x_2. (You may assume $x_1 < x_2$).

Fit the simple linear model to the data, and show that

$$\hat{b} = \frac{\bar{y}_2 - \bar{y}_1}{x_2 - x_1}.$$

where $\bar{y}_1$ is the average of $y_1, \ldots, y_{10}$ and $\bar{y}_2$ is the average of $y_{11}, \ldots, y_{20}$

8. A test was administered to a sample of California high school students. The sample was stratified according to location of the schools. Make a 90% confidence interval for the proportion of high school students that would pass the exam.

	population	sample size	number passed
Inner City	20 mil	100	25
Suburban	40 mil	400	135
Rural	30 mil	900	250