Final Exam: The final exam will be **Tuesday, December 18, from 8:00 P.M. to 11:00 P.M.**, in the usual classroom. It will be cumulative, covering the entire semester. The exam will be open book: you may use the text for the class and your notes. However, calculators may not be used.

These problems concentrate on material we have covered **since the last exam.** To review earlier material, use the review problems distributed for the first two exams, as well as the exams themselves (and, of course, homework problems).

There will be a problem session Sunday, December 16, in the usual classroom, starting at 4:00 P.M., to discuss these problems, earlier problems, and any other questions you have. You can also ask me questions via email.

1. (a) A random sample $X_1, \ldots, X_n$ is drawn from a binomial distribution with parameters θ and k; k is known but θ is not. Find the maximum likelihood estimator of θ, and the estimator obtained from the method of moments. (In each case the answer is $\bar{X}/k$.)

(b) It is known that a certain quantity takes values 0, 1, 2, and 3. One hundred independent measurements of this quantity yield these values 5, 26, 33, and 36 times, respectively. Test, at the 5% level of significance, the hypothesis that the quantity is distributed with binomial distribution $b(x; 3, \theta)$ for some θ.

2. Two random variables X and Y have joint density

$$f(x,y) = \begin{cases} C(2 + y - xy), & \text{if } 0 \le x \le 2 \text{ and } 0 \le y \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the constant C.

(b) Show that the regression $E(X \mid y)$ is linear, and find the regression coefficients α and β.

(c) Find the regression $E(Y \mid x)$ and show that it is not linear.

3. Two random variables X and Y have joint density

$$f(x,y) = Ke^{-2(x^2+3xy+4y^2)/7}.$$

Explain why this is a bivariate normal distribution and, by comparing the given density with the standard form of the bivariate normal density, find $E(X)$, $E(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X,Y)$, and the constant K. (Hint: $\rho = -3/4$, but first try to do the problem without using this information.)

4. Suppose that X and Y have bivariate normal distribution, with parameters $\mu_1 = \mu_2 = 0$, $\sigma_1 = 3$, $\sigma_2 = 2$, and $\rho = -\sqrt{3}/2$. Find

(a) The probability that $0 \le Y \le 2$.

(b) The probability that $0 \le Y \le 2$, given that $X = \sqrt{3}$.

(c) Find the mean and variance of the random variable $2X - 3Y$.

5. Work out Exercise 14.53.

6. (a) Using the data from Exercise 14.57, construct a 99% confidence interval for the regression coefficient β, and test the null hypothesis $\beta \le 0.5$ at the 5% level of significance;

(b) Assuming the data of (a) comes from a bivariate normal distribution, find a 95% confidence interval for the correlation coefficient ρ.

7. (a) Write down a 3×3 table containing integers between 10 and 50.

(b) Make up a word problem for which the table from (a) represents the data, and in which it is necessary to test whether the population attribute described by the rows is independent of the attribute described by the columns. Carry out the test.

(c) Make up a word problem for which the table from (a) represents the data, and in which it is necessary to test whether the population attribute described by the columns has the same distribution in the various populations described by the rows. Carry out the test.