

STAT 613 Midterm (1 hour 15 minutes) April 13th, 2012

Marks will be given for clarity of the solution.

Good Luck!

- (1) The object of this question is to use the log-likelihood ratio test to derive the χ -squared test for independence (in the case of two by two tables). In other words, derive the distribution of the test statistic

$$T = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} + \frac{(O_3 - E_3)^2}{E_3} + \frac{(O_4 - E_4)^2}{E_4},$$

under the null that there is no association between the categorical variables C and R , where and $E_1 = n_3 \times n_1/N$, $E_2 = n_4 \times n_1/N$, $E_3 = n_3 \times n_2/N$ and $E_4 = n_4 \times n_2/N$.

	C_1	C_2	Subtotal
R_1	O_1	O_2	n_1
R_2	O_3	O_4	n_2
Subtotal	n_3	n_4	N

State all results you use. (hint: It may be useful to use the Taylor approximation $x \log(x/y) \approx (x - y) + \frac{1}{2}(x - y)^2/y$). [10]

- (2) Consider the following shifted exponential mixture distribution

$$f(x; \lambda_1, \lambda_2, p, a) = p \frac{1}{\lambda_1} \exp(-x/\lambda_1) I(x \geq 0) + (1 - p) \frac{1}{\lambda_2} \exp(-(x - a)/\lambda_2) I(x \geq a),$$

where p, λ_1, λ_2 and a are unknown.

[15]

- (i) Make a plot of the above mixture density.

Considering the cases $x \geq a$ and $x < a$ separately, calculate the probability of belonging to each of the mixtures, given the observation X_i (ie. Define the variable δ_i , where $P(\delta_i = 0) = p$, $f(x|\delta_i = 0) = \frac{1}{\lambda_1} \exp(-x/\lambda_1)$ etc. and calculate $P(\delta_i = 0|X_i = x)$ and $P(\delta_i = 1|X_i = x)$).

- (ii) Show how the EM-algorithm can be used to estimate $a, p, \lambda_1, \lambda_2$. At each iteration you should be able to obtain explicit solutions for *most* of the parameters, give as many details as you can.

Hint: It may be beneficial for you to use profiling too.

- (iii) From your knowledge of estimation of these parameters, what do you conjecture the rates of convergence to be? Will they all be the same, or possibly different?

- (iv) Not part of the exams: code the estimator. Through simulations try to verify your conjecture in (iii).