

# STAT 830

## Problems: Assignment 3

1. Suppose  $\{X_{ij}; j = 1, \dots, n_i; i = 1, \dots, p\}$  are independent  $N(\mu_i, \sigma^2)$  random variables. (This is the usual set-up for the one-way layout.)
  - (a) Find the MLE's for  $\mu_i$  and  $\sigma$ .
  - (b) Find the expectations and variances of these estimators.
2. Let  $T_i$  be the error sum of squares in the  $i$ th cell in the previous question. In this question you may assume that

$$\frac{T_i}{\sigma^2} = \frac{\sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2}{\sigma^2}$$

has a  $\chi^2$  distribution with  $n_i - 1$  degrees of freedom.

- (a) Find the best estimate of  $\sigma^2$  of the form  $\sum_1^p a_i T_i$  in the sense of mean squared error. That is, find the constants  $a_i$  which minimize the mean squared error.
  - (b) Do the same under the condition that the estimator must be unbiased.
  - (c) Find the joint density of the  $T_i$ .
  - (d) If only  $T_1, \dots, T_p$  are observed what is the MLE of  $\sigma$ ? It is a fact which you may assume that the  $T_i$  are independent.
3. In question 1 take  $n_i = 2$  for all  $i$  and let  $p \rightarrow \infty$ . What is the limit of the MLE of  $\sigma$ ? Hint: find the mean and variance of the mle of  $\sigma^2$ .
  4. Suppose that  $Y_1, \dots, Y_n$  are independent random variables and that  $x_1, \dots, x_n$  are the corresponding values of some covariate. Suppose that the density of  $Y_i$  is

$$f(y_i) = \exp(-y_i \exp(-\alpha - \beta x_i) - \alpha - \beta x_i) 1(y_i > 0)$$

where  $\alpha$ , and  $\beta$  are unknown parameters.

- (a) Find the log-likelihood, the score function and the Fisher information.
- (b) For the following data set fit the model and produce a contour plot of the log-likelihood surface, the profile likelihood for  $\beta$  and an approximate 95% confidence interval for  $\beta$ .

$x_i$	$Y_i$	$x_i$	$Y_i$
-0.59	0.694	0.31	0.009
0.28	0.408	-0.45	0.032
-1.08	1.330	1.23	0.057
0.79	0.266	-1.71	1.530
-0.15	0.111	0.43	0.152
0.50	0.382	-0.61	0.358
0.54	0.165	0.60	0.214
-1.05	2.766	0.68	0.375
0.66	0.103	0.92	0.193
0.40	0.309	1.50	0.245

5. Consider the random effects one way layout. You have data  $X_{ij}; i = 1, \dots, p; j = 1, \dots, n$  and a model  $X_{ij} = \mu + \alpha_i + \epsilon_{ij}$  where the  $\alpha$ 's are iid  $N(0, \tau^2)$  and the  $\epsilon$ 's are iid  $N(0, \sigma^2)$ . The  $\alpha$ s are independent of the  $\epsilon$ s.
- (a) Compute the mean and variance covariance matrix of the vector you get by writing out all the  $X_{ij}$  as a vector.
- (b) Suppose that  $M$  is a matrix of the form  $aI + b11^t$  where  $I$  is a  $p \times p$  identity and  $1$  denotes a column vector of  $p$  ones. Show that  $M^{-1}$  is of the form  $cI + d11^t$  and find  $c$  and  $d$ . In what follows you may use the fact that the determinant of  $M$  is  $a^{p-1}(a + pb)$ .
- (c) Write down the likelihood.
- (d) Find the MLEs.
6. For each of the doses  $d_1, \dots, d_p$  a number of animals  $n_1, \dots, n_p$  are treated with the corresponding dose of some drug. The number dying at dose  $d$  is Binomial with parameter  $h(d)$ . A common model for  $h(d)$  is  $\log\{h/(1-h)\} = \alpha + \beta d$ .
- (a) Find the likelihood equations for estimating  $\alpha$  and  $\beta$ .

- (b) Find the Fisher information matrix.
- (c) Define the parameter LD50 as the value of  $d$  for which  $h(d) = 1/2$ ; express LD50 as a function of  $\alpha$  and  $\beta$ .
- (d) Use a Taylor expansion to find large sample confidence limits for LD50.
- (e) At each of the doses -3.204, -2.903, -2.602, -2.301 and -2.000 a sample of 40 mice were exposed to antipneumonococcus serum. The numbers surviving were 7, 18, 32, 35, and 38 respectively. Get numerical values for the theory above. You can use *glm* or get preliminary estimates based on linear regression of the MLE of  $h(d_i)$  against dose.

Q33

7. Suppose  $X_1, \dots, X_n$  are a sample of size  $n$  from the density

$$f_{\alpha,\beta}(x) = \frac{1}{\beta\Gamma(\alpha)} \left(\frac{x}{\beta}\right)^{\alpha-1} \exp(-x/\beta) 1(x > 0).$$

In the following question the digamma function  $\psi$  is defined by  $\psi(\alpha) = \frac{d}{d\alpha} \log(\Gamma(\alpha))$  and the trigamma function  $\psi'$  is the derivative of the digamma function. From the identity  $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$  you can deduce recurrence relations for the digamma and trigamma functions.

- (a) For  $\alpha = \alpha_o$  known find the mle for  $\beta$ .
- (b) When both  $\alpha$  and  $\beta$  are unknown what equation must be solved to find  $\hat{\alpha}$ , the mle of  $\alpha$ ?
- (c) Evaluate the Fisher information matrix.
- (d) A sample of size 40 is given below:

3.547	1.228	2.052	1.556	2.487	0.469	2.707	0.395
0.770	0.666	4.242	1.474	1.277	2.519	0.578	2.989
1.900	1.422	3.701	1.278	2.820	0.224	0.482	1.426
2.146	2.975	2.792	0.846	3.190	1.680	0.686	1.634
0.969	4.010	1.792	1.287	0.730	0.849	2.447	2.147

Use this data in the following questions. First take  $\alpha = 1$  and find the mle of  $\beta$  subject to this restriction.

- (e) Now use  $E(X) = \alpha\beta$  and  $\text{Var}(X) = \alpha\beta^2$  to get method of moments estimates  $\tilde{\alpha}$  and  $\tilde{\beta}$  for the parameters.
- (f) Do two steps of Newton Raphson to get MLEs.
- (g) Use Fisher's scoring idea, which is to replace the second derivative in Newton Raphson with the Fisher information (and then not change it as you run the iteration), to redo the previous question.
- (h) Compute standard errors for the MLEs and compare the difference between the estimates in the previous 2 questions to the SEs.

Due: TBA