Inference Tutorial 6

1. Consider the *linear mixed model*

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2 + b_j + \epsilon_{ij} \text{ where } b_j \underset{\text{ind}}{\sim} N(0, \sigma_b^2) \text{ and } \epsilon_{ij} \underset{\text{ind}}{\sim} N(0, \sigma^2) \tag{1}$$

(a) The standard generalized likelihood ratio test can not be used to test $H_0 : \sigma_b = 0$. Explain what the problem is with reference to the derivation of the large sample distribution of the GLRT test statistic, or the conditions given in section 8.8.

(b) Does the same problem apply to testing $H_0 : \beta_2 = 0$ in this model? Explain your answer.

(c) An alternative model to the one above is that the structure is the same, except that the residuals sharing the same j are autocorrelated, such that the autocorrelation between ϵ_{ij} and $\epsilon_{i'j}$ is $\rho^{|i-i'|}$, where $-1 < \rho < 1$, and ρ is a parameter to be estimated. The alternative models are estimated by MLE and compared using a GLRT. Explain whether or not this comparison is valid.

(d) Suppose that you are working as a consultant statistician and a scientiest comes to you with the following problem. They have fitted model (1) to some data and the maximum likelihood estimate $\hat{\sigma}_b = 0$. The scientist wants to test $H_0 : \sigma_b = 0$, but knows that this is problematic. What would you advise them?

Solution

(a) The log likelihood ratio statistic asymptotic distribution relies on the maximum likelihood estimators having a normal distribution centred on the true parameter value, in the large sample limit (see final paragraph of 8.6), but such large sample normality can not possibly occur when the estimated parameter is always $\geq$ the true parameter value (see section 8.4 — the required derivative of the log likelihood at the true parameter value is not even properly defined).

(b) No. $\beta_2 = 0$ is not on the edge of the feasible parameter space. There is no problem.

(c) The test is fine. The simpler model is equivalent to setting $\rho = 0$, which is not at the edge of the feasible parameter space, so no problem.

(d) If your best estimate of a parameter corresponds exactly to the null hypothesis of interest then no sane test is ever going to reject the null. It doesn't matter what the exact distribution of the likelihood ratio statistics is: its observed value here is the *smallest value it could possibly take*, so the p-value must be 1.

2. Random variables X and Y have joint p.d.f.

$$f(x, y) = kx^\alpha y^\beta \quad 0 \leq x \leq 1, \;\; 0 \leq y \leq 1$$

Assume that you have n independent pairs of observations (x_i, y_i).

(a) Evaluate k in terms of the α and β.

(b) Find the maximum likelihood estimators of α and β.

(c) Find the approximate variances of your estimators.

Solution

(a)

$$k \int_0^1 \int_0^1 x^\alpha y^\beta dx dy = \frac{k}{(\alpha+1)(\beta+1)} = 1 \Rightarrow k = (\alpha+1)(\beta+1)$$

(b)

$$l(\alpha, \beta) = n\log(\alpha+1) + n\log(\beta+1) + \sum \alpha \log(x_i) + \sum \beta \log(y_i)$$

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha+1} + \sum \log(x_i) \qquad \frac{\partial l}{\partial \beta} = \frac{n}{\beta+1} + \sum \log(y_i)$$

Setting both these to zero and solving yields:

$$\hat{\alpha} = \frac{-n}{\sum \log(x_i)} - 1 \quad \hat{\beta} = \frac{-n}{\sum \log(y_i)} - 1$$

(c)

$$\frac{\partial^2 l}{\partial\alpha\partial\beta} = 0 \quad \frac{\partial^2 l}{\partial\beta^2} = \frac{-n}{(\beta+1)^2} \quad \frac{\partial^2 l}{\partial\alpha^2} = \frac{-n}{(\alpha+1)^2}$$

$$\Rightarrow \text{var}(\hat\alpha) \simeq (\hat\alpha+1)^2/n \quad \text{var}(\hat\beta) \simeq (\hat\beta+1)^2/n.$$

3. This question continues from question 2 of sheet 5 (the urchin model), and concerns confidence intervals and hypothesis tests.

   (a) Adding the argument `hessian=TRUE` to a call to `optim` or `nlm` will cause the functions to return a numerical approximation to the second derivative (Hessian) matrix of the objective function (negative log likelihood, in our case). Use this fact and an appropriate general large sample result to obtain 95% confidence intervals for the parameters of the urchin model.

   (b) Produce a modified version of your log likelihood function in which you can fix the value of ϕ at a numerical value, rather than supplying it as part of the parameter vector. Using this, test $H_0 : \phi = 1$ with a GLRT.

   (c) Use your function from the last part to find a 90% CI for ϕ (This uses test inversion as covered earlier, or in section 8.7).

Solution

(a)

```
fit <- optim(theta0,urchin.ll,method="BFGS",hessian=TRUE,vol=uv$vol,age=uv$age)
theta.hat <- fit$par
cov.th <- solve(fit$hessian)
theta.sd <- diag(cov.th)^.5
exp(rbind(theta.hat - 1.96*theta.sd,theta.hat + 1.96*theta.sd))
             [,1]     [,2]      [,3]      [,4]
[1,] 8.003984e-05 1.093921 0.2199385 0.4196200
[2,] 4.173154e+00 1.288381 2.9656426 0.5295134
```

(b)

```
urchin.ll0 <- function(theta,age,vol,phi) {
  theta <- exp(theta)
  omega <- theta[1]
  gamma <- theta[2]
  sigma <- theta[3]
  am <- log(phi/(gamma*omega))/gamma ## critical age
  ind <- age < am      ## before critical age?
  V <- age*0 ## create expected volume vector
  V[ind] <- omega*exp(gamma*age[ind])
  V[!ind] <- phi/gamma + phi*(age[!ind]-am)
  ll <- sum(dnorm(sqrt(vol),sqrt(V),sigma,log=TRUE)) ## log lik.
  attr(ll,"V") <- V
  -ll ## return -ve computed log likelihood
} ## urchin.ll

fit <- optim(theta0,urchin.ll,method="BFGS",vol=uv$vol,age=uv$age)
fit0 <- optim(theta0[-2],urchin.ll0,method="BFGS",vol=uv$vol,age=uv$age,phi=1)
lambda <- 2*(fit0$value - fit$value) ## ’value’ is negative log lik
pchisq(lambda,df=1,lower.tail=FALSE) ## quite good evidence against
[1] 1.500281e-05
```

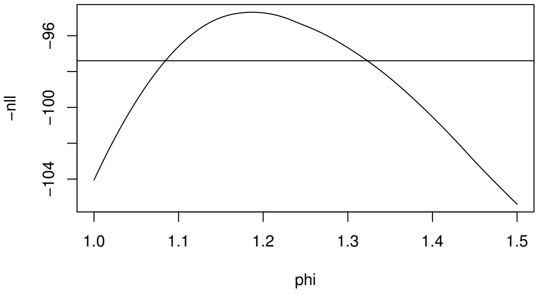
. . . firm evidence against the null hypothesis.

(c)

```
dlcrit <- qchisq(.9,df=1)/2 ## threshold log-lik difference for accept reject at 10%
nll <- phi <- 1 + 0:1000/2000
for (i in 1:length(phi)) { ## loop over phi values, getting log lik for each
  nll[i] <- optim(fit$par[-2],urchin.ll0,method="BFGS",vol=uv$vol,age=uv$age,
                phi=phi[i])$value
}
range(phi[-nll>-fit$value-dlcrit])
[1] 1.113 1.277
```

You might also want to produce the following plot, illustrating the log likelihood and the accept/reject threshold

```
plot(phi,-nll,type="l")
abline(-fit$value-dlcrit,0)
```


−nll
phi