APTS module *Statistical Inference*

January 2011

Assessment material

The work provided here is intended to take students up to half a week to complete. Students should talk to their supervisors to find out whether their department requires this work as part of any formal accreditation process (APTS itself has no resources to assess or certify students). It is anticipated that departments will decide on the appropriate *level* of assessment locally, and may choose to drop some (or indeed all) of the items, accordingly.

Students should also take their supervisor's advice on how much total time to devote to these problems. A reasonable target would be to do at least two of the questions from Section A and one from Section B.

Section A: consolidation of the APTS-week material

1. [From Part 2. If you get stuck, any good book on the analysis of binary/categorical data should have some discussion of this, e.g., Cox & Snell (1989) *Analysis of Binary Data*.] For the binary matched pairs model, derive the conditional binomial distribution for inference on the common log odds ratio ψ. Discuss whether it is reasonable to discard all the data from 'non-mixed' pairs.

2. [From Part 3.] Let $Y_1, \ldots, Y_n$ have independent Poisson distributions with mean μ. Obtain the maximum likelihood estimator of μ. Obtain that estimator's variance,

   (a) from first principles;

   (b) by the general results of asymptotic theory.

   Suppose now that it is observed only whether each observation is zero or non-zero.

   (c) What now are the maximum likelihood estimate of μ and its asymptotic variance?

   (d) At what value of μ is the ratio of the latter to the former variance minimized?

   (e) In what practical context might these results be relevant?

3. [From Part 3.] Suppose that $Y_1, \ldots, Y_n$ are independent, with $Y_i \sim N(\lambda + \psi x_i, \sigma^2)$ and σ^2 known.

   (a) Calculate the expected information matrix $i(\psi, \lambda)$, and relate this to what you know about least squares.

   (b) Find a new parameterization, (ψ, τ) say, in which τ is orthogonal to ψ. What are the advantages of orthogonality?

Section B: extension of the APTS-week material

1. [Extends Part 4.] Let $Y_1, \ldots, Y_n$ be independently binomially distributed each corresponding to ν trials with probability of success θ. Both ν and θ are unknown. Construct simple (inefficient) estimates of the parameters, for example by considering

   - the mean and variance of the sample
   - or the proportions of values equal to zero and one

   On the basis of one or both of these preliminary estimates for what combinations of ν, θ would you expect the maximum likelihood estimate of ν to be at infinity with appreciable probability? Simulate one of these situations and either

   - examine the shape of the likelihood surface for say 10 simulation runs

   or (more advanced)

   - study how the proportion of formally infinite estimates depends on the underlying parameters
   - or study the properties of profile-likelihood based confidence limits in such a situation
   - or set up a Bayesian formulation.

   The situation described is a simplified version of a model for the estimation of the number of bugs in a complex piece of computer software. (e.g., Joe and Reid, 1985, *JASA* **80**, 222–226)

2. [Extends Part 5.] Write a short summary (2 pages or so) of the uses and limitations of *empirical Bayes* methods. [Not covered in this week's APTS lectures; see for example the books by Cox (2006; sec 5.12), Davison (2003; sec 11.5) or O'Hagan and Forster (2004; sec 5.25–5.27) for some discussion and further references.]