

Problem 1 Suppose we hope to eliminate the nuisance parameters λ from a likelihood $L(\psi, \lambda)$ by using an integrated likelihood

$$\int L(\psi, \lambda) d\lambda.$$

- (a) Criticise this approach by computing the integrated likelihoods when the likelihood is based on two independent exponential variables with parameters (i) λ and $\lambda\psi$, (ii) $1/\lambda$ and ψ/λ , where $\lambda, \psi > 0$.
- (b) Now suppose that in (i) the parameters are given a density $\pi(\psi, \lambda)$ and we compute the resulting marginal density for ψ . Show that if the corresponding prior density is used in the parametrization in (ii), the problems in (a) do not arise.

Problem 2 A random sample $y_1, \dots, y_n$ of distinct observations has arisen from an unknown distribution function G . Consider a multinomial distribution in which $p_j = P(Y = y_j)$, for $j = 1, \dots, n$.

- (a) Use Lagrange multipliers to show that the empirical distribution function with $p_j \equiv 1/n$ maximises the likelihood $\sum_j \log p_j$ of the observed data subject to the constraints $p_j \geq 0$ and $\sum_{j=1}^n p_j \leq 1$.
- (b) Now add the constraint that $\sum_j p_j c_j(\theta) = 0$, where $c_j(\theta) \equiv c(y_j; \theta)$ is a $d \times 1$ function of y_j and θ ; this is the empirical version of the constraint $E\{c(Y; \theta)\} = 0$, with expectation taken over $Y \sim G$. Show that in this case the log likelihood for a specific θ is the *empirical likelihood*

$$\ell_E(\theta) = \sum_{j=1}^n \log\{1 + \lambda^T c_j(\theta)\}, \quad \text{where } \lambda \equiv \lambda_\theta \text{ satisfies } 0 = \sum_{j=1}^n \frac{c_j(\theta)}{1 + \lambda^T c_j(\theta)}.$$

- (c) If $c_j(\theta) = y_j - \theta$, show that the maximum empirical likelihood estimate is $\hat{\theta}_E = \bar{y}$, with $\lambda = 0$, and that in general the p_j satisfy $\sum y_j p_j = \theta$, where $\min y_j < \theta < \max y_j$. Does this make sense to you?

Problem 3 In testing $\alpha \neq 1$ when $Y_1, \dots, Y_n$ is a random sample from the gamma (α, λ) distribution, with λ unknown (Problem 1(b) of Week 10), show that (a) the distribution of the test statistic is invariant to λ , and (b) that λ may be eliminated by appropriate conditioning. In case (a) say how you would simulate the distribution of the test statistic. Does the same algorithm apply in (b)?

Problem 4

- (a) If $P \sim U(0, 1)$, show that $Q = 2 \min(P, 1 - P) \sim U(0, 1)$.
- (b) What are the achievable significance levels for testing that a single Poisson variable has mean $\mu_0 = 2$, with alternative mean (i) greater than 2 and (ii) less than 2? Does this matter for computing confidence intervals?

Problem 5 Is the test used in the top quark example most powerful? Against which alternatives?