

MATH562 – Fall 2025

Problem Set: Week 11

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 $\lambda, \psi > 0$.

Solution: In case (i) the original likelihood is $\lambda e^{-\lambda y_1} \times \lambda\psi e^{\lambda\psi y_2}$, so the integrated likelihood is

$$\int_0^\infty \psi \lambda^2 \exp\{-\lambda(y_1 + \psi y_2)\} \, d\lambda = \frac{\psi}{(y_1 + \psi y_2)^3}, \int_0^\infty u^2 e^{-u} \, du = \frac{2\psi}{(y_1 + \psi y_2)^3}, \quad \psi > 0,$$

where we set $u = \lambda(y_1 + \psi y_2)$. In case (ii) the original likelihood is $\lambda^{-1} e^{-y_1/\lambda} \times (\psi/\lambda) e^{\psi y_2/\lambda}$, giving

$$\int_0^\infty \psi \lambda^{-2} \exp\{-(y_1 + \psi y_2)/\lambda\} \, d\lambda = \frac{\psi}{y_1 + \psi y_2} \int_0^\infty u^2 e^{-u} u^{-2} \, du = \frac{\psi}{y_1 + \psi y_2}, \quad \psi > 0,$$

where we set $u = (y_1 + \psi y_2)/\lambda$. Hence the result depends on the nuisance parametrisation, which is clearly unsatisfactory.

- (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. **Solution:** In case (i) we now obtain

$$\int L(\psi, \lambda) \pi(\psi, \lambda) \, d\lambda = \int_0^\infty \psi \lambda^2 \exp\{-\lambda(y_1 + \psi y_2)\} \pi(\psi, \lambda) \, d\lambda,$$

but the prior corresponding to $(\psi, \mu = 1/\lambda)$ is $\pi^*(\psi, \mu) = \pi(\psi, 1/\mu)\mu^{-2}$, giving

$$\int_0^\infty L(\psi, 1/\mu) \pi^*(\psi, \mu) \, d\mu = \int_0^\infty \psi \mu^{-2} \exp\{-(y_1 + \psi y_2)/\mu\} \pi(\psi, 1/\mu) \mu^{-2} \, d\mu,$$

and this reduces to the result for (i) when we change the variable of integration to $\lambda = 1/\mu$.

- *2. A random sample y_1, \dots, y_n of distinct observations, $y_i \neq y_j$ for $i \neq j$, has arisen from an unknown distribution function G . Consider a multinomial distribution in which $p_j = \Pr(Y = y_j)$, 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 $\sum_{j=1}^n p_j \leq 1$.

Hint: first argue that it's sufficient to consider the case where $p_j > 0$ for all j and $\sum_{j=1}^n p_j = 1$.

- (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 , maximizing over the p_j given the constraints, is the *empirical log-likelihood*

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

- (c) If $c_j(\theta) = y_j - \theta$, show that the maximum empirical likelihood estimate is $\hat{\theta}_E = \bar{y}$, with $\lambda_\theta = 0$, and that for any valid θ we have $\sum y_j p_j = \theta$ and $\min_j y_j < \theta < \max_j y_j$.

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(c) of Week 10),

- (a) show that the distribution of the test statistic is invariant to λ
- (b) describe how you would simulate the distribution of the test statistic in (a) to obtain p -values without relying on the asymptotic distribution for $n \rightarrow \infty$.
- (c) show that λ may be eliminated by appropriate conditioning.

Solution:

(a) Under the null hypothesis we can write $Y_j \stackrel{d}{=} E_j/\lambda$, where $E_1, \dots, E_n \text{ i.i.d. exp}(1)$, so the test statistic

$$T = \sum_{j=1}^n \log(Y_j/\bar{Y}) \stackrel{d}{=} \sum_{j=1}^n \log(E_j/\bar{E}),$$

which does not depend on λ . Hence the statistic is invariant to λ .

- (b) This could be simulated by generating $E_1, \dots, E_n \text{ i.i.d. exp}(1)$ and hence computing a null distribution for T and its quantiles.
- (c) Under the null hypothesis the data are exponential, so the minimal sufficient statistic is $S = Y_1 + \dots + Y_n$, which has a gamma (n, λ) distribution. Hence the conditional density of the data given S is

$$\frac{\lambda^n \exp\{-\lambda(y_1 + \dots + y_n)\}}{\lambda^n s^{n-1} \exp(-\lambda s)/\Gamma(n)} = \frac{\Gamma(n)}{s^{n-1}}, \quad 0 < y_1, \dots, y_n < s, \sum y_j = s.$$

This is the uniform distribution on an n -dimensional simplex, and of course it does not depend on λ . The statistic T depends only on the $Y_j/\bar{Y} = nY_j/S$, which is invariant to s (so in particular the same simulation algorithm as in (a) will work).

4. (a) If $P \sim U(0, 1)$, show that $Q = 2 \min(P, 1 - P) \sim U(0, 1)$. **Solution:** For $x \in (0, 1)$ and because the events $P \leq u$ and $P \geq 1 - u$ are disjoint for $u < 1/2$ we have

$$\Pr(Q \leq x) = \Pr\{\min(P, 1 - P) \leq x/2\} = \Pr(P \leq x/2) + \Pr(P \geq 1 - x/2) = x/2 + x/2 = x,$$

as required.

- (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? **Solution:** These are computed using the R code The values are fairly limited in both cases, though the limitations for (ii) are not surprising.

```
> ppois(1:7,lambda=2,lower.tail=FALSE)
[1] 0.593994150 0.323323584 0.142876540 0.052653017 0.016563608 0.004533806 0.001096719
> ppois(0:2,lambda=2)
[1] 0.1353353 0.4060058 0.6766764
```

For confidence intervals we would solve the equation $\Pr(Y \geq 2; \mu) = \alpha$ for some specific values of α , so the discreteness is not a major issue.