ST2132 Tutorial 8 2210

1. Let $x_1, \ldots, x_n$ be realisations of IID Normal(μ, σ^2) random variables $X_1, \ldots, X_n$. The parameter of interest is $\theta = (\mu, \sigma)$.

   (a) Write down the logarithm of the normal density at x_i. Hence write down the loglikelihood function $\ell(\theta)$, with the refinement if you like.

   (b) Show that the solutions of $\frac{\mathrm{d}\ell}{\mathrm{d}\theta} = 0$ is $(\bar{x}, \hat{\sigma})$. Write down $\bar{x}$ and $\hat{\sigma}$ in terms of the data.

   (c) Show that

$$\frac{\partial^2 \ell}{\partial \mu^2} = -\frac{n}{\sigma^2}, \qquad \frac{\partial^2 \ell}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{3}{\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\frac{\partial^2 \ell}{\partial \mu \partial \sigma} = \frac{\partial^2 \ell}{\partial \sigma \partial \mu} = -\frac{2n}{\sigma^3}(\bar{x} - \mu)$$

   Hence show that at $(\bar{x}, \hat{\sigma})$ the determinant of $\frac{\mathrm{d}^2\ell}{\mathrm{d}\theta^2}$ is positive.

   (d) Let $f(x, y)$ be a function of two variables. The point satisfying $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$ is a maximum if at that point, $\frac{\partial^2 f}{\partial x^2} < 0$, $\frac{\partial^2 f}{\partial y^2} < 0$, and $\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \frac{\partial^2 f}{\partial x \partial y}\frac{\partial^2 f}{\partial y \partial x} > 0$. Verify that this implies $(\bar{x}, \hat{\sigma})$ are the ML estimates of (μ, σ).

2. Let $x_1, \ldots, x_n$ be realisations of IID Normal(μ, ν) RV's $X_1, \ldots, X_n$. $\nu > 0$ is the variance.

   (a) Show that the loglikelihood is $\ell(\mu, \nu) = -\frac{n}{2}\log \nu - \frac{1}{2\nu}\sum_{i=1}^{n}(x_i - \mu)^2$.

   (b) Show that

$$\frac{\partial^2 \ell}{\partial \mu^2} = -\frac{n}{\nu}, \qquad \frac{\partial^2 \ell}{\partial \nu^2} = \frac{n}{2\nu^2} - \frac{1}{\nu^3}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\frac{\partial^2 \ell}{\partial \mu \partial \nu} = \frac{\partial^2 \ell}{\partial \nu \partial \mu} = -\frac{n}{\nu^2}(\bar{x} - \mu)$$

   Find the ML estimates of (μ, ν), justifying your answer with the general fact in 1(d).

   (c) We usually write $\nu = \sigma^2$, with $\sigma > 0$. What can you say about the ML estimates of ν and σ?

3. A die was rolled 60 times independently and under the same conditions. The results are summarised in the table. You are encouraged to use vectorised operations in R to work on this problem.

Value	1	2	3	4	5	6
Frequency	4	6	17	16	8	9

(a) Set up a statistical model for the data, in terms of random variables and their joint distribution.

(b) Use MOM to estimate $\mathbf{p} = (p_1, \ldots, p_6)$ to two decimal places, where p_i is the probability that the rolled die shows i.

(c) What are the exact SE's of the estimates?

(d) Calculate approximate SE's by the bootstrap, to three decimal places.

(e) By looking at intervals of the form (estimate − 1.96SE, estimate + 1.96SE), do the data suggest that the die is unfair?

4. A corn plant can have starchy or sugary kernels, and leaves with green or white base. In a certain breeding experiment, new plants are created independently and under identical conditions. In 1927, Carver observed 3839 such new plants, which are classified as follows.

Type	*Count*
Starchy green	1997
Starchy white	906
Sugary green	904
Sugary white	32
Total	3839

By genetic theory, the type probabilities are

Type	*Chance*	*Chance in* θ
Starchy green	$(r^2 - 2r + 3)/4$	$(2 + \theta)/4$
Starchy white	$(2r - r^2)/4$	$(1 - \theta)/4$
Sugary green	$(2r - r^2)/4$	$(1 - \theta)/4$
Sugary white	$(1 - r)^2/4$	$\theta/4$

Table 1: $1 - r$ is the recombination fraction between the two loci. $\theta = (1 - r)^2$.

(a) Set up a statistical model for the data, in terms of random variables and their joint distribution, with parameter θ.

(b) Given that $1/2 \leq r \leq 1$, what is the parameter space of θ?

(c) Show that the ML estimate of θ satisfies the equation

$$3839x^2 + 1655x - 64 = 0$$

Hence obtain the ML estimate to two decimal places.