

Chapter 3

General results for multi-parameter problems

In this chapter we will study some general results for multi-parameter problems.

3.1 Different levels of prior knowledge

We have *substantial prior information* for θ when the prior distribution *dominates* the posterior distribution, that is $\pi(\theta|\mathbf{x}) \sim \pi(\theta)$.

When prior information about θ is limited, this is usually represented through the use of a conjugate prior distribution, with vague prior knowledge represented by making the conjugate distribution as diffuse as possible.

If we represent prior ignorance for a single parameter θ by using uniform or improper priors then we have seen (MAS2903, section 3.4) that, in general, the prior for $g(\theta)$ is not constant and so we are not ignorant about $g(\theta)$. The same problem occurs when we have more than one parameter.

Suppose we represent prior ignorance about $\theta = (\theta_1, \theta_2, \dots, \theta_p)^T$ using $\pi(\theta) = \text{constant}$. Let $\phi_i = g_i(\theta)$, $i = 1, \dots, p$ and $\phi = (\phi_1, \dots, \phi_p)^T$ be a 1-1 transformation. Then, in general, the prior density for ϕ is not constant and this suggests that we are not ignorant about ϕ . However, if we are ignorant about θ then we must also be ignorant about $g(\theta)$.

This contradiction makes it impossible to use this representation of prior ignorance.

Example 3.1

Suppose $0 < \theta_1 < 1$ and $0 < \theta_2 < 1$. If we are ignorant about $\boldsymbol{\theta} = (\theta_1, \theta_2)^T$ then show that $\theta_1\theta_2$ does not have a constant prior density.



Example 3.2

Suppose we have a random sample from a $N(\mu, 1/\tau)$ distribution (with τ unknown). Determine the Jeffreys prior for this model.

Hint: We have already seen that the likelihood function can be written as

$$f(\mathbf{x}|\mu, \tau) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left[-\frac{n\tau}{2} \{s^2 + (\bar{x} - \mu)^2\}\right]$$

where

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$





...Solution to Example 3.2



...Solution to Example 3.2

3.2 Asymptotic posterior distribution

Suppose we have a statistical model for data with likelihood function $f(\mathbf{x}|\boldsymbol{\theta})$, where $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ and $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)^T$, together with a prior distribution with density $\pi(\boldsymbol{\theta})$ for $\boldsymbol{\theta}$. Then

$$J(\hat{\boldsymbol{\theta}})^{1/2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})|\mathbf{x} \xrightarrow{\mathcal{D}} N_p(0, I_p) \quad \text{as } n \rightarrow \infty,$$

where $\hat{\boldsymbol{\theta}}$ is the likelihood mode, I_p is the $p \times p$ identity matrix and $J(\boldsymbol{\theta})$ is the observed information matrix, with $(i, j)^{th}$ element

$$J_{ij} = -\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f(\mathbf{x}|\boldsymbol{\theta}),$$

and $A^{1/2}$ denotes the square root matrix of A .

Comments

Example 3.3

Suppose we now have a random sample from a $N(\mu, 1/\tau)$ distribution (with unknown precision). Determine the asymptotic posterior distribution for (μ, τ) .

Hint: we have already seen that the likelihood function can be written as

$$f(\mathbf{x}|\mu, \tau) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left[-\frac{n\tau}{2} \{s^2 + (\bar{x} - \mu)^2\}\right]$$

where $s^2 = \sum_{i=1}^n (x_i - \bar{x})^2/n$.





...Solution to Example 3.3

3.3 Learning objectives

By the end of this chapter, you should be able to:

- understand different levels of prior information and have an appreciation for the difficulty of specifying ignorance priors in multi-parameter problems
- determine the asymptotic posterior distribution when the data are a large random sample from **any** distribution
- explain the similarities and differences between the asymptotic posterior distribution and the asymptotic distribution of the maximum likelihood estimator

