On Independent Reference Priors

Mi Hyun Lee

Dissertation submitted to the faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Statistics

Dongchu Sun, Chairman

Eric P. Smith, Co-Chairman

George R. Terrell

Pang Du

December 5, 2007

Blacksburg, Virginia

Keywords: Independent reference priors, Reference priors, Probability matching priors

On Independent Reference Priors

Mi Hyun Lee

(ABSTRACT)

In Bayesian inference, the choice of prior has been of great interest. Subjective priors are ideal if sufficient information on priors is available. However, in practice, we cannot collect enough information on priors. Then objective priors are a good substitute for subjective priors.

In this dissertation, an independent reference prior based on a class of objective priors is examined. It is a reference prior derived by assuming that the parameters are independent. The independent reference prior introduced by Sun and Berger (1998) is extended and generalized. We provide an iterative algorithm to derive the general independent reference prior. We also propose a sufficient condition under which a closed form of the independent reference prior is derived without going through the iterations in the iterative algorithm. The independent reference prior is then shown to be useful in respect of the invariance and the first order matching property. It is proven that the independent reference prior is invariant under a type of one-to-one transformation of the parameters. It is also seen that the independent reference prior is a first order probability matching prior under a sufficient condition. We derive the independent reference priors for various examples. It is observed that they are first order matching priors and the reference priors in most of the examples. We also study an independent reference prior in some types of non-regular cases considered by Ghosal (1997).

This dissertation is dedicated to the memory of

my father in heaven

Acknowledgments

This dissertation would never have been possible without the expert guidance of my committee members, and support from my family and friends.

I wish to express my sincere gratitude to my esteemed advisor, Dr. Dongchu Sun, for persevering with me throughout the time it took me to complete this dissertation. I would also like to thank Dr. Eric P. Smith, who advised and supported me as my co-advisor. My thanks go out to Dr. George R. Terrell and Dr. Pang Du for their contribution.

Many people on the faculty and staff of the department of Statistics assisted and encouraged me in various ways. I also thank them.

I am grateful for encouragement and support from my mother and late father. I must acknowledge as well my friends, who were always cheering me up.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

In Bayesian inference, the selection of prior has been of great interest and various kinds of priors have been proposed. There are two categories of priors based on the amount of information on priors that we could have, which are *subjective priors* and *objective priors* (or *noninformative priors*). If sufficient information on priors is available, subjective priors could be a good choice. Unfortunately, in practice, we might not often collect enough information. Then noninformative priors or objective priors, which are derived only by using the assumed model and the available data, can be used as a substitute for subjective priors. Thus the use of noninformative or objective priors has increased in Bayesian analysis.

Many kinds of noninformative priors have been developed: constant priors [Laplace (1812)], Jeffreys priors [Jeffreys (1961)], reference priors [Bernardo (1979), Berger and Bernardo (1992)], independent reference priors [Sun and Berger (1998)], probability matching priors [Datta and Mukerjee (2004)], and noninformative priors in non-regular cases

[Ghosal and Samanta (1997), Ghosal (1997), Ghosal (1999)]. We review them precisely in Section 1.2.

We study an *independent reference prior* which originated in Sun and Berger (1998). It is a reference prior derived with the assumption of the independence of the parameters. In many practical problems, we can obtain partial information on priors such as the independence of the parameters. Then independent reference priors could be used for such situations.

In this dissertation, the independent reference prior introduced by Sun and Berger (1998) is extended and generalized. We consider multiple groups of parameters while Sun and Berger (1998) used two groups of parameters. An iterative algorithm to compute the general independent reference prior is proposed. Then a mild sufficient condition to make an inference on the result of the iterative algorithm without going through the iterations is also provided. The independent reference prior holds the invariance and the first order matching property. We prove that our independent reference prior is invariant under a type of one-to-one reparameterization where the Jacobian matrix is diagonal. A sufficient condition under which the independent reference prior is a first order matching prior is given. Then the independent reference priors are derived in numerous examples. It turns out that they are first matching priors and the reference priors in most of the examples. Additionally, we present an iterative algorithm to obtain an independent reference prior in some types of non-regular cases where the support of the data is either monotonically increasing or decreasing in a non-regular type parameter. It is verified that the independent reference prior is a first order matching prior under a sufficient condition. Some examples are also given.

1.2 Literature Review

The history of objective priors is described in this section. Constant priors [Section 1.2.1], Jeffreys priors [Section 1.2.2], reference priors [Section 1.2.3], independent reference priors [Section 1.2.4], Probability matching priors [Section 1.2.5], and objective priors in non-regular cases [Section 1.2.6] are reviewed.

1.2.1 Constant Priors

Objective priors began with a constant prior (or a flat prior) which is just proportional to 1. Laplace (1812) employed it for Bayesian analysis. The constant prior is very simple and easy to use. However it is not invariant to one-to-one transformations of the parameters.

1.2.2 Jeffreys Priors

Jeffreys (1961) proposed a rule for deriving a prior which is invariant to any one-to-one reparameterization. It is called a Jeffreys-rule prior which is still one of the popular objective priors. The Jeffreys-rule prior is proportional to the positive square root of the determinant of the Fisher information matrix defined as (1.1). The Fisher information is a measure of the amount of information about the parameters, provided by the data from model. Datta and Ghosh (1996) pointed out that the Jeffreys-rule prior performs satisfactorily in one-parameter cases but poorly in multi-parameter cases. An inconsistent Bayes estimator or an unreasonable posterior were produced in some of multi-parameter examples. Thus the use of the Jeffreys-rule prior is somewhat controversial in multi-parameter cases. Jeffreys (1961) recommended an independence Jeffreys prior which could modify the deficiencies of the Jeffreys-rule prior in multi-parameter cases. It is the product of the Jeffreys-rule priors for each group of parameters when the other groups of parameters are held fixed.

1.2.3 Reference Priors

Bernardo (1979) introduced a reference prior which fixes the deficiencies of the Jeffreys-rule prior in multi-parameter problems. The ad hoc modifications which are required for the Jeffreys-rule prior in multi-parameter situations are not necessary for the reference prior. Bernardo (1979) separated the parameters into the parameters of interest and nuisance parameters, and considered the parameters sequentially in the process of deriving a reference prior. Then a reference prior is more successful in multi-parameter cases. A reference prior is defined as a prior which maximizes asymptotically the expected information provided by the data from model about the parameters, which is the same as the expected Kullback-Leibler divergence between the posterior and prior. Then the reference prior has minimal influence since the data has maximal influence on the inference. Bernardo (1979) just introduced the basic idea of reference priors and posteriors without the mathematical details for their construction.

The idea of Bernardo (1979) was broadened and generalized by Berger and Bernardo (1992). They divided the parameters into two or more groups according to their order of inferential importance. They provided an in-depth description of mathematical methods to derive a reference prior.

Now the reference prior method is described in detail. Let us start with the notation that is necessary to explain the method. Consider a parametric family of distributions whose density is given by $f(x;\boldsymbol{\theta})$ for the data $X \in \mathcal{X}$, where $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^p$ is a p-dimensional unknown parameter vector which can be decomposed into m sub-groups

$$\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m).$$

Here $\boldsymbol{\theta}_i = (\theta_{i1}, \ldots, \theta_{i,p_i}) \in \boldsymbol{\Theta}_i \subset \mathbb{R}^{p_i}$, $\boldsymbol{\Theta} = \boldsymbol{\Theta}_1 \times \cdots \times \boldsymbol{\Theta}_m$ with $p_1 + \cdots + p_m = p$.

We define the Fisher information matrix of $\boldsymbol{\theta}$

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = -E_{\boldsymbol{\theta}}\left[\frac{\partial^2}{\partial\boldsymbol{\theta}_i\partial\boldsymbol{\theta}_j}\log f(X;\boldsymbol{\theta})\right], \quad i,j = 1, \ldots, m, \tag{1.1}$$

where $E_{\boldsymbol{\theta}}$ denotes expectation over X given $\boldsymbol{\theta}$. We will often write $\boldsymbol{\Sigma}$ instead of $\boldsymbol{\Sigma}(\boldsymbol{\theta})$.

Also define, for $j = 1, \ldots, m$,

$$\begin{aligned} \boldsymbol{\theta}_{[j]} &= (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_j), \\ \boldsymbol{\theta}_{[\sim j]} &= (\boldsymbol{\theta}_{j+1}, \ldots, \boldsymbol{\theta}_m), \end{aligned}$$

where $\boldsymbol{\theta}_{[\sim 0]} = \boldsymbol{\theta}$ and $\boldsymbol{\theta}_{[0]}$ is vacuous.

Let $Z_t = \{X_1, \ldots, X_t\}$ be the random variable that would arise from t conditionally independent replications of the original experiment. Then Z_t has density

$$p(z_t|\boldsymbol{\theta}) = \prod_{i=1}^{t} f(x_i;\boldsymbol{\theta}). \tag{1.2}$$

First, we see how to develop a reference prior for regular cases in the sense that $p(z_t \mid \boldsymbol{\theta})$, given by (1.2), is asymptotically normally distributed. Assume that $\boldsymbol{\Sigma}$ is invertible and let $\boldsymbol{S} = \boldsymbol{\Sigma}^{-1}$. Write $\boldsymbol{S}$ as

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{21}^t & \cdots & \boldsymbol{A}_{m1}^t \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{m2}^t \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}_{m1} & \boldsymbol{A}_{m2} & \cdots & \boldsymbol{A}_{mm} \end{pmatrix}$$

so that $\boldsymbol{A}_{ij}$ is $p_i \times p_j$, and define $\boldsymbol{S}_j$ to be the upper left $\left(\sum_{k=1}^{j} p_k\right) \times \left(\sum_{k=1}^{j} p_k\right)$ corner of $\boldsymbol{S}$ with $\boldsymbol{S}_m \equiv \boldsymbol{S}$ and $\boldsymbol{H}_j \equiv \boldsymbol{S}_j^{-1}$. Then the matrices $\boldsymbol{h}_j$, defined to be the lower right $p_j \times p_j$ corner of $\boldsymbol{H}_j$, $j = 1, \ldots, m$, will be of central importance. Note that $\boldsymbol{h}_1 \equiv \boldsymbol{H}_1 \equiv \boldsymbol{A}_{11}^{-1}$ and if $\boldsymbol{S}$ is a block diagonal matrix, that is $\boldsymbol{A}_{ij} = 0$ for all $i \neq j$, then $\boldsymbol{h}_j \equiv \boldsymbol{A}_{jj}^{-1}$, $j = 1, \ldots, m$. Finally, if $\boldsymbol{\Theta}^* \subset \boldsymbol{\Theta}$, we define

$$\boldsymbol{\Theta}^*(\boldsymbol{\theta}_{[j]}) = \{\boldsymbol{\theta}_{j+1} : (\boldsymbol{\theta}_{[j]}, \boldsymbol{\theta}_{j+1}, \boldsymbol{\theta}_{[\sim j+1]}) \in \boldsymbol{\Theta}^* \text{ for some } \boldsymbol{\theta}_{[\sim j+1]}\}.$$

$|\boldsymbol{A}|$ denotes the determinant of $\boldsymbol{A}$, and $\mathbf{1}_{\boldsymbol{\Omega}}(y)$ equals 1 if $y \in \boldsymbol{\Omega}$, 0 otherwise.

The reference prior method for regular cases can be described in four steps.

1. Choose a nested sequence $\boldsymbol{\Theta}^1 \subset \boldsymbol{\Theta}^2 \subset \cdots$ of compact subsets of $\boldsymbol{\Theta}$ such that $\bigcup_{l=1}^{\infty} \boldsymbol{\Theta}^l = \boldsymbol{\Theta}$. This step is not necessary if the reference priors turn out to be proper.

2. Order the coordinates $(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m)$. Usually, the order should typically be according to inferential importance; in particular, the first group of parameters should be of interest. Note that $(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m)$ is assumed to be ordered for convenience of notation.

3. To start, define

$$\begin{aligned} \pi_m^l(\boldsymbol{\theta}_{[\sim m-1]}|\boldsymbol{\theta}_{[m-1]}) &= \pi_m^l(\boldsymbol{\theta}_m|\boldsymbol{\theta}_{[m-1]}) \\ &= \frac{|\boldsymbol{h}_m(\boldsymbol{\theta})|^{1/2}\mathbf{1}_{\boldsymbol{\Theta}^l(\boldsymbol{\theta}_{[m-1]})}(\boldsymbol{\theta}_m)}{\int_{\boldsymbol{\Theta}^l(\boldsymbol{\theta}_{[m-1]})} |\boldsymbol{h}_m(\boldsymbol{\theta})|^{1/2} d\boldsymbol{\theta}_m}. \end{aligned}$$

For $j = m-1, \ldots, 1$, define

$$\pi_j^l(\boldsymbol{\theta}_{[\sim j-1]}|\boldsymbol{\theta}_{[j-1]}) = \frac{\pi_{j+1}^l(\boldsymbol{\theta}_{[\sim j]}|\boldsymbol{\theta}_{[j]}) \exp\left\{\frac{1}{2}E_j^l[(\log|\boldsymbol{h}_j(\boldsymbol{\theta})|)|\boldsymbol{\theta}_{[j]}]\right\} \mathbf{1}_{\boldsymbol{\Theta}^l(\boldsymbol{\theta}_{[j-1]})}(\boldsymbol{\theta}_j)}{\int_{\boldsymbol{\Theta}^l(\boldsymbol{\theta}_{[j-1]})} \exp\left\{\frac{1}{2}E_j^l[(\log|\boldsymbol{h}_j(\boldsymbol{\theta})|)|\boldsymbol{\theta}_{[j]}]\right\} d\boldsymbol{\theta}_j},$$

where

$$E_j^l[g(\boldsymbol{\theta})|\boldsymbol{\theta}_{[j]}] = \int_{\{\boldsymbol{\theta}_{[\sim j]}:(\boldsymbol{\theta}_{[j]},\boldsymbol{\theta}_{[\sim j]})\in\boldsymbol{\Theta}^l\}} g(\boldsymbol{\theta})\pi_{j+1}^l(\boldsymbol{\theta}_{[\sim j]}|\boldsymbol{\theta}_{[j]})d\boldsymbol{\theta}_{[\sim j]}.$$

For $j = 1$, write

$$\pi_1^l(\boldsymbol{\theta}) = \pi_1^l(\boldsymbol{\theta}_{[\sim 0]}|\boldsymbol{\theta}_{[0]}).$$

4. Define a reference prior, $\pi(\boldsymbol{\theta})$, as any prior for which

$$E_l^X D(\pi_1^l(\boldsymbol{\theta}|X), \pi(\boldsymbol{\theta}|X)) \to 0 \quad \text{as} \quad l \to \infty,$$

where the Kullback-Leibler divergence between two densities g and h on $\boldsymbol{\Theta}$ is denoted by

$$D(g, h) = \int_{\boldsymbol{\Theta}} g(\boldsymbol{\theta}) \log \left[\frac{g(\boldsymbol{\theta})}{h(\boldsymbol{\theta})}\right] d\boldsymbol{\theta},$$

and E_l^X is expectation with respect to

$$p^l(x) = \int_{\boldsymbol{\Theta}} f(x; \boldsymbol{\theta}) \pi_1^l(\boldsymbol{\theta}) d\boldsymbol{\theta}.$$

Typically, $\pi(\boldsymbol{\theta})$ is determined by the simple relation

$$\pi(\boldsymbol{\theta}) = \lim_{l \to \infty} \frac{\pi_1^l(\boldsymbol{\theta})}{\pi_1^l(\boldsymbol{\theta}^*)},$$

where $\boldsymbol{\theta}^*$ is an interior point of $\boldsymbol{\Theta}$.

Definitely, a reference prior depends on the grouping and the ordering of the parameters. Thus Berger and Bernardo (1992) recommended deriving a reference prior by considering one parameter per group in Step 2. We call such a reference prior a one-at-a-time reference prior. However one-at-a-time reference priors still depend on the order of inferential importance of the parameters. Note that it can be easily shown that a reference prior is equivalent to the Jeffreys-rule prior in one-parameter cases.

Datta and Ghosh (1996) provided another expression for $|\boldsymbol{h}_j(\boldsymbol{\theta})|$, $j = 1, \ldots, m$. Write the Fisher information matrix of $\boldsymbol{\theta}$ in partitioned form as

$$\boldsymbol{\Sigma} = \Big((\boldsymbol{\Sigma}_{ij})\Big), \quad i, j = 1, \ldots, m.$$

Also write for $j = 0, \ldots, m-1$,

$$\boldsymbol{\Sigma}_{[\sim jj]} = \Big((\boldsymbol{\Sigma}_{ik})\Big), \quad i, k = j+1, \ldots, m.$$

Then

$$|\boldsymbol{h}_j(\boldsymbol{\theta})| = \frac{|\boldsymbol{\Sigma}_{[\sim j-1, j-1]}|}{|\boldsymbol{\Sigma}_{[\sim jj]}|}, \quad j = 1, \ldots, m,$$

where $|\boldsymbol{\Sigma}_{[\sim mm]}| = 1$.

Next, the reference prior method for non-regular cases which was proposed by Berger and Bernardo (1992) is shown. Only Step 3 is different from the regular cases. Thus we just describe Step 3.

3$'$. For $j = m, m-1, \ldots, 1$, iteratively compute densities

$$\pi_j^l(\boldsymbol{\theta}_{[\sim j-1]}|\boldsymbol{\theta}_{[j-1]}) \propto \pi_{j+1}^l(\boldsymbol{\theta}_{[\sim j]}|\boldsymbol{\theta}_{[j]}) h_j^l(\boldsymbol{\theta}_j|\boldsymbol{\theta}_{[j-1]}),$$

where $\pi_{m+1}^l \equiv 1$ and h_j^l is computed by the following two steps.

3$'$a: Define $p_t(\boldsymbol{\theta}_j|\boldsymbol{\theta}_{[j-1]})$ as

$$p_t(\boldsymbol{\theta}_j|\boldsymbol{\theta}_{[j-1]}) \propto \exp\left\{\int p(z_t|\boldsymbol{\theta}_{[j]}) \log p(\boldsymbol{\theta}_j|z_t, \boldsymbol{\theta}_{[j-1]}) dz_t\right\}, \tag{1.3}$$

where

$$\begin{aligned} p(z_t|\boldsymbol{\theta}_{[j]}) &= \int p(z_t|\boldsymbol{\theta}) \pi_{j+1}^l(\boldsymbol{\theta}_{[\sim j]}|\boldsymbol{\theta}_{[j]}) d\boldsymbol{\theta}_{[\sim j]}, \\ p(\boldsymbol{\theta}_j|z_t, \boldsymbol{\theta}_{[j-1]}) &\propto p(z_t|\boldsymbol{\theta}_{[j]}) p_t(\boldsymbol{\theta}_j|\boldsymbol{\theta}_{[j-1]}). \end{aligned}$$

3$'$b: Assuming the limit exists, define

$$h_j^l(\boldsymbol{\theta}_j|\boldsymbol{\theta}_{[j-1]}) = \lim_{t\to\infty} p_t(\boldsymbol{\theta}_j|\boldsymbol{\theta}_{[j-1]}). \tag{1.4}$$

For $j = 1$, write

$$\pi_1^l(\boldsymbol{\theta}) = \pi_1^l(\boldsymbol{\theta}_{[\sim 0]}|\boldsymbol{\theta}_{[0]}).$$

Berger and Bernardo (1992) pointed out that in practice it is very hard to compute the p_t, given by (1.3), and to find their limit in (1.4).

Berger (1992) [attributed to Ghosh and Mukerjee (1992)] introduced a reverse reference prior, which is obtained by reversing the roles of the interest parameters and nuisance parameters when deriving a reference prior, in order to satisfy the probability matching criterion when the parameters are orthogonal. We explain the probability matching criterion in Section 1.2.5.

According to Datta and Ghosh (1996), the invariance of the reference prior is valid under a type of one-to-one reparameterization where the Jacobian matrix is upper triangular. However the reverse reference prior does not remain invariant to any one-to-one reparameterization. Datta and M. Ghosh (1995) compared reference priors and reverse reference priors. They provided a sufficient condition under which the two priors agree.

1.2.4 Independent Reference Priors

Sun and Berger (1998) derived conditional reference priors when partial information is available. They considered three situations. When a subjective conditional prior density is given, two methods to find a marginal reference prior were described. When a subjective marginal prior is known, a conditional reference prior was proposed. When two groups of parameters are assumed to be independent, an independent reference prior was defined.

The independent reference prior is our main focus in this dissertation. In most examples of Bayesian inference, the reference priors are expressed as the product of marginal reference priors. If we have information on the independence of the groups of parameters, we can surely use an independent reference prior which does not depend on the order of inferential importance of the groups of parameters instead of a reference prior which depends on it.

Assuming the independence of two groups of parameters, $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, Sun and Berger (1998) suggested the following iterative algorithm to derive an independent reference prior. Note that $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ is the Fisher information matrix of $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$, $\boldsymbol{\Sigma}_{22} = \boldsymbol{\Sigma}_{22}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ is the Fisher information matrix of $\boldsymbol{\theta}_2$, given that $\boldsymbol{\theta}_1$ is held fixed, and $\boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{11}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ is the Fisher information matrix of $\boldsymbol{\theta}_1$, given that $\boldsymbol{\theta}_2$ is held fixed.

Algorithm A:

Step 0. Choose any initial nonzero marginal prior density for $\boldsymbol{\theta}_2$, $\pi_2^{(0)}(\boldsymbol{\theta}_2)$, say.

Step 1. Define an interim prior density for $\boldsymbol{\theta}_1$ by

$$\pi_1^{(1)}(\boldsymbol{\theta}_1) \propto \exp\left\{\frac{1}{2}\int \pi_2^{(0)}(\boldsymbol{\theta}_2)\log\left(\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{22}|}\right) d\boldsymbol{\theta}_2\right\}.$$

Step 2. Define an interim prior density for $\boldsymbol{\theta}_2$ by

$$\pi_2^{(1)}(\boldsymbol{\theta}_2) \propto \exp\left\{\frac{1}{2}\int \pi_1^{(1)}(\boldsymbol{\theta}_1)\log\left(\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}|}\right) d\boldsymbol{\theta}_1\right\}.$$

Replace $\pi_2^{(0)}$ in *Step 0* by $\pi_2^{(1)}$ and repeat *Step 1* and *2* to obtain $\pi_1^{(2)}$ and $\pi_2^{(2)}$. Consequently, we generate two sequences $\{\pi_1^{(l)}\}_{l\geq 1}$ and $\{\pi_2^{(l)}\}_{l\geq 1}$. The desired marginal reference priors will be the limits

$$\pi_i^R = \lim_{l\to\infty} \pi_i^{(l)}, \quad i = 1, 2,$$

if the limits exist.

Sun and Berger (1998) pointed out that in applying the iterative algorithm, it may be necessary to operate on compact sets, and then let the sets grow as the reference prior method. They also established a sufficient condition under which the result of the algorithm is inferred without going through the iterations.

1.2.5 Probability Matching Priors

The concept of probability matching priors are quite different from the previous objective priors. Welch and Peers (1963) introduced the basic idea of probability matching priors. Datta and Mukerjee (2004) summarized and discussed various probability matching priors. The priors satisfying the criterion that the frequentist coverage probabilities of Bayesian credible sets agree asymptotically to the Bayesian coverage probabilities of the credible sets up to a certain order, are defined as probability matching priors or simply matching priors. In other words, the difference between the frequentist confidence sets and the Bayesian credible sets should be small in an asymptotic way. There are several matching criteria. For example, the matching might be carried out through posterior quantiles, distribution functions, highest posterior density regions, inversion of certain test statistics, or prediction. For each matching criterion, the differential equations which matching priors must satisfy were derived.

Matching priors for posterior quantiles are most popular. First and second order matching priors are widely used for posterior quantile matching priors. We consider only a first order matching prior in this dissertation. The differential equation which a first order matching prior must satisfy was introduced by Datta and J. K. Ghosh (1995) and revisited by Datta and Mukerjee (2004). Referring to Chapter 2 of Datta and Mukerjee (2004), matching priors for posterior quantiles are defined as follows. Consider priors $\pi(\cdot)$ for which the relation

$$P_{\boldsymbol{\theta}}\{\psi \leq \psi^{(1-\alpha)}(\pi, X)\} = 1 - \alpha + o(n^{-r/2}), \tag{1.5}$$

holds for $r = 1, 2, \ldots$, and for each $\alpha \in (0, 1)$. n is the sample size, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$, where $\theta_i \in \boldsymbol{\Theta}_i \subset \mathbb{R}$ is an unknown parameter vector, $\psi = \psi(\boldsymbol{\theta})$ is the one-dimensional parametric function of interest, $P_{\boldsymbol{\theta}}\{\cdot\}$ is the frequentist probability measure under $\boldsymbol{\theta}$, and $\psi^{(1-\alpha)}(\pi, X)$

is the $(1-\alpha)^{th}$ posterior quantile of ψ under $\pi(\cdot)$, given the data X. Then priors satisfying (1.5) for $r = 1$ are called first order matching priors for ψ. First order matching priors π^M for ψ must satisfy the following differential equation,

$$\frac{\partial}{\partial\theta_1}(\xi_1\pi) + \cdots + \frac{\partial}{\partial\theta_m}(\xi_m\pi) = 0, \tag{1.6}$$

where

$$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)' = \frac{\boldsymbol{\Sigma}^{-1}\nabla\psi}{\sqrt{\nabla\psi'\boldsymbol{\Sigma}^{-1}\nabla\psi}} \tag{1.7}$$

with $\nabla\psi = \left(\frac{\partial}{\partial\theta_1}\psi, \ldots, \frac{\partial}{\partial\theta_m}\psi\right)'$ and $\boldsymbol{\Sigma}$ is the Fisher information matrix of $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$.

By Welch and Peers (1963), the Jeffreys-rule prior is a first order matching prior in one-dimensional parameter cases. Thus a reference prior is also a first order matching prior in one-parameter cases. Remember that a reverse reference prior was introduced by Berger (1992) to meet the matching criterion under orthogonality. Datta and M. Ghosh (1995) reaffirmed that a reverse reference prior must be a matching prior under orthogonal parameterizations but a reference prior does not need to be even under orthogonality. By Datta and Ghosh (1996), a probability matching prior was shown to be invariant under any one-to-one reparameterization.

1.2.6 Non-regular Cases

The concept and algorithm for reference priors for non-regular cases were proposed by Bernardo (1979) and Berger and Bernardo (1992). It is however hard to apply in practice. See Section 1.2.3 for details.

Ghosal and Samanta (1997) obtained a reference prior in one-parameter non-regular cases such as a one-parameter family of discontinuous densities where the support of the data is either monotonically decreasing or increasing in the parameter. They derived a reference prior by maximizing the expected Kullback-Leibler divergence between the prior and the posterior in an asymptotic way.

Ghosal (1997) proposed a reference prior in multi-parameter non-regular cases such as a multi-parameter family of discontinuous densities where some regular type parameters are added to the one-parameter family of discontinuous densities used by Ghosal and Samanta (1997). The reference prior was computed through two procedures when nuisance parameter exists. One procedure adapted the reference prior method proposed by Berger and Bernardo (1992) and another was an extension of the reference prior method provided by Ghosal and Samanta (1997).

The differential equations which first order matching priors for one- and multi-parameter non-regular cases must satisfy were built by Ghosal (1999). He considered the one- and multi-parameter family of discontinuous densities used by Ghosal and Samanta (1997) and Ghosal (1997).

1.3 Outline

This dissertation is organized as follows.

In Chapter 2, a general independent reference prior is developed by extending the results of Sun and Berger (1998). The invariance under a type of one-to-one transformation of the parameters is proven. The first order matching property is also obtained.

The independent reference priors are derived in various examples of probability distributions in Chapter 3. We compare the independent reference priors with the reference priors.

We also see whether the independent reference priors satisfy the first order matching criterion or not.

In Chapter 4, an independent reference prior in some types of non-regular cases is derived. We construct a sufficient condition under which the independent reference prior agrees to a first order matching prior. The independent reference priors are computed in some examples.

We summarize and propose future work in Chapter 5.

Chapter 2

Main Results for Independent Reference Priors

2.1 Notation

Consider a parametric family of distributions whose density is given by $f(x;\boldsymbol{\theta})$ for the data $X \in \mathcal{X}$, where $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^p$ is a p-dimensional unknown parameter vector which can be decomposed into m sub-vectors

$$\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m). \tag{2.1}$$

Here $\boldsymbol{\theta}_i = (\theta_{i1}, \ldots, \theta_{i,p_i}) \in \boldsymbol{\Theta}_i \subset \mathbb{R}^{p_i}$, where $\boldsymbol{\Theta} = \boldsymbol{\Theta}_1 \times \cdots \times \boldsymbol{\Theta}_m$ with $p_1 + \cdots + p_m = p$.

We define the partitioned Fisher information matrix of $\boldsymbol{\theta}$

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = -E_{\boldsymbol{\theta}}\left[\frac{\partial^2}{\partial\boldsymbol{\theta}_i\partial\boldsymbol{\theta}_j}\log f(X;\boldsymbol{\theta})\right], \quad i,j = 1, \ldots, m, \tag{2.2}$$

where $E_{\boldsymbol{\theta}}$ denotes expectation over X given $\boldsymbol{\theta}$. We will often write $\boldsymbol{\Sigma}$ instead of $\boldsymbol{\Sigma}(\boldsymbol{\theta})$.

2.2 Independent Reference Priors

In this section, we provide an independent reference prior by generalizing the results of Sun and Berger (1998). We consider more groups of parameters than the two groups considered by Sun and Berger (1998). We propose an iterative algorithm to find the general independent reference prior and obtain a mild sufficient condition to deduce the result of it without going through the iterations. Thus a closed form of the independent reference prior is derived. The invariance of independent reference priors to a type of one-to-one reparameterization where the Jacobian matrix is diagonal is proven. A sufficient condition under which the independent reference prior agrees to a first order matching prior is proposed. Thus two desired figures of independent reference priors are achieved. Numerous examples are given in Chapter 3. We study an independent reference prior in some types of non-regular cases in Chapter 4.

Now we present an iterative algorithm to derive an independent reference prior for $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m)$. It is an extension of *Algorithm A* proposed by Sun and Berger (1998). To begin with, we note that $\boldsymbol{\Sigma}_{ii}^c$ is the matrix obtained by removing the rows and columns corresponding to $\boldsymbol{\theta}_i$ from $\boldsymbol{\Sigma}$, and $\boldsymbol{\theta}_i^c = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_{i+1}, \ldots, \boldsymbol{\theta}_m)$, $i = 1, \ldots, m$.

Algorithm B:

Step 0. Choose any initial nonzero marginal prior densities for $\boldsymbol{\theta}_i$, namely $\pi_i^{(0)}(\boldsymbol{\theta}_i)$ for all $i = 2, \ldots, m$.

Step 1. Define an interim prior density for $\boldsymbol{\theta}_1$ by

$$\pi_1^{(1)}(\boldsymbol{\theta}_1) \propto \exp\left\{\frac{1}{2}\int \prod_{k=2}^{m} \pi_k^{(0)}(\boldsymbol{\theta}_k) \log\left(\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}^c|}\right) d\boldsymbol{\theta}_1^c\right\}.$$

Step i. For $i = 2, \cdots, m$, define interim prior densities for $\boldsymbol{\theta}_i$ by

$$\pi_i^{(1)}(\boldsymbol{\theta}_i) \propto \exp\left\{\frac{1}{2}\int \left[\prod_{k=1}^{i-1} \pi_k^{(1)}(\boldsymbol{\theta}_k)\right]\left[\prod_{k=i+1}^{m} \pi_k^{(0)}(\boldsymbol{\theta}_k)\right] \log\left(\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{ii}^c|}\right) d\boldsymbol{\theta}_i^c\right\}$$

Replace $\pi_i^{(0)}(\boldsymbol{\theta}_i)$, $i = 2, \ldots, m$, in *Step 0* by $\pi_i^{(1)}(\boldsymbol{\theta}_i)$, $i = 2, \ldots, m$, and repeat *Step i* to obtain $\pi_i^{(2)}(\boldsymbol{\theta}_i)$ for $i = 1, \ldots, m$. Consequently, the sequences of the marginal priors $\{\pi_i^{(l)}(\boldsymbol{\theta}_i) : i = 1, \ldots, m\}_{l \geq 1}$ are generated. The marginal reference priors for $\boldsymbol{\theta}_i$ will be the limits

$$\pi_i^R(\boldsymbol{\theta}_i) = \lim_{l \to \infty} \pi_i^{(l)}(\boldsymbol{\theta}_i), \quad i = 1, \ldots, m,$$

if the limits exist.

In practice, the interim priors $\{\pi_i^{(l)}(\boldsymbol{\theta}_i) : i = 1, \ldots, m\}_{l \geq 1}$ could be improper. In such cases, one might need to implement an algorithm using compact sets as it is recommended by Sun and Berger (1998). Choose an increasing sequence $\{\boldsymbol{\Theta}_i^j\}_{j \geq 1}$ of compact subsets of $\boldsymbol{\Theta}_i$ such that

$$\bigcup_{j=1}^{\infty} \boldsymbol{\Theta}_i^j = \boldsymbol{\Theta}_i, \quad i = 1, \ldots, m.$$

We then could use the following algorithm.

Algorithm B′:

Step 0. For fixed j, choose any initial proper marginal prior densities for $\boldsymbol{\theta}_i$ on $\boldsymbol{\Theta}_i^j$, namely $\pi_{ij}^{(0)}(\boldsymbol{\theta}_i)$ for all $i = 2, \ldots, m$.

Step 1. Define an interim prior density for $\boldsymbol{\theta}_1$ on $\boldsymbol{\Theta}_1^j$ by

$$\pi_{1j}^{(1)}(\boldsymbol{\theta}_1) \propto \exp\left\{\frac{1}{2}\int_{\otimes_{h=2}^m \boldsymbol{\Theta}_h^j} \prod_{k=2}^{m} \pi_{kj}^{(0)}(\boldsymbol{\theta}_k) \log\left(\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}^c|}\right) d\boldsymbol{\theta}_1^c\right\}.$$

Step i. For $i = 2, \ldots, m$, define interim prior densities for $\boldsymbol{\theta}_i$ on $\boldsymbol{\Theta}_i^j$ by

$$\pi_{ij}^{(1)}(\boldsymbol{\theta}_i) \propto \exp\left\{\frac{1}{2}\int_{\otimes_{h\neq i}\boldsymbol{\Theta}_h^j}\left[\prod_{k=1}^{i-1}\pi_{kj}^{(1)}(\boldsymbol{\theta}_k)\right]\left[\prod_{k=i+1}^{m}\pi_{kj}^{(0)}(\boldsymbol{\theta}_k)\right]\log\left(\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{ii}^c|}\right)d\boldsymbol{\theta}_i^c\right\}.$$

Replace $\pi_{ij}^{(0)}(\boldsymbol{\theta}_i)$, $i = 2, \ldots, m$, in *Step 0* by $\pi_{ij}^{(1)}(\boldsymbol{\theta}_i)$, $i = 2, \ldots, m$, and repeat *Step i* to obtain $\pi_{ij}^{(2)}(\boldsymbol{\theta}_i)$ for $i = 1, \ldots, m$. Consequently, we have sequences of marginal priors $\{\pi_{ij}^{(l)}(\boldsymbol{\theta}_i) : i = 1, \ldots, m\}_{j\geq 1, l\geq 1}$. Let $\boldsymbol{\theta}_i^0$ be an interior point of $\boldsymbol{\Theta}_i$, $i = 1, \ldots, m$. The marginal reference priors for $\boldsymbol{\theta}_i$ will be the limits

$$\pi_i^R(\boldsymbol{\theta}_i) = \lim_{j\to\infty}\lim_{l\to\infty}\frac{\pi_{ij}^{(l)}(\boldsymbol{\theta}_i)}{\pi_{ij}^{(l)}(\boldsymbol{\theta}_i^0)}, \quad i = 1, \ldots, m,$$

if these limits exist.

The convergence of the iterations is not guaranteed. Then we might not obtain a closed form of the independent reference prior. Here we have found a sufficient condition for deriving an independent reference prior without going through the iterations.

Theorem 2.1 *Suppose, for all $i = 1, \ldots, m$,*

$$\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{ii}^c|} = f_{1i}(\boldsymbol{\theta}_i)f_{2i}(\boldsymbol{\theta}_i^c), \tag{2.3}$$

where $\boldsymbol{\theta}_i^c = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_{i+1}, \ldots, \boldsymbol{\theta}_m)$, $\boldsymbol{\Sigma}$ is the Fisher information matrix of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m)$, and $\boldsymbol{\Sigma}_{ii}^c$ is the matrix which is derived by removing the rows and columns corresponding to $\boldsymbol{\theta}_i$ from $\boldsymbol{\Sigma}$. Then the independent reference prior for $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m)$ is

$$\pi^R(\boldsymbol{\theta}) = \prod_{i=1}^{m}\pi_i^R(\boldsymbol{\theta}_i), \tag{2.4}$$

where the marginal reference priors for $\boldsymbol{\theta}_i$, $i = 1, \ldots, m$, are

$$\pi_i^R(\boldsymbol{\theta}_i) \propto \sqrt{f_{1i}(\boldsymbol{\theta}_i)}. \tag{2.5}$$

Proof. It can be easily seen that under Condition (2.3), $\pi_i^R(\boldsymbol{\theta}_i)$, $i = 1, \ldots, m$, do not depend on the choices of $\pi_i^{(0)}(\boldsymbol{\theta}_i)$, $i = 2, \ldots, m$, in *Step 0.* Hence the marginal reference priors for $\boldsymbol{\theta}_i$, $i = 1, \ldots, m$, have the form of (2.5) and the independent reference prior for $\boldsymbol{\theta}$ is given by (2.4). □

In the next corollary an independent reference prior is derived under orthogonality. Consequently it is shown to be same as the the independent reference prior in (2.4).

Corollary 2.1 *Consider the following Fisher information matrix of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m)$,*

$$\boldsymbol{\Sigma} = diag\Big(f_{11}(\boldsymbol{\theta}_1)f_{21}(\boldsymbol{\theta}_1^c), \ldots, f_{1m}(\boldsymbol{\theta}_m)f_{2m}(\boldsymbol{\theta}_m^c)\Big).$$

Then the independent reference prior for $\boldsymbol{\theta}$ is the same as (2.4).

Proof. It is clear that for all $i = 1, \ldots, m$, $|\boldsymbol{\Sigma}|/|\boldsymbol{\Sigma}_{ii}^c| = f_{1i}(\boldsymbol{\theta}_i)f_{2i}(\boldsymbol{\theta}_i^c)$, which satisfies Condition (2.3). □

Now we prove that the independent reference prior, given by (2.4), is invariant under a type of one-to-one transformation of the parameters where the Jacobian matrix is diagonal.

Theorem 2.2 *For any $i = 1, \cdots, m$, let $\boldsymbol{\eta}_i = g_i(\boldsymbol{\theta}_i)$ be a one-to-one transformation of $\boldsymbol{\theta}_i$. Then under Condition (2.3), the independent reference prior for $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, \cdots, \boldsymbol{\eta}_m)$ is formed as*

$$\pi^R(\boldsymbol{\eta}) = \prod_{i=1}^m \pi_i^R(\boldsymbol{\eta}_i), \tag{2.6}$$

where the marginal reference priors for $\boldsymbol{\eta}_i$, $i = 1, \ldots, m$, are

$$\pi_i^R(\boldsymbol{\eta}_i) \propto \sqrt{f_{1i}\Big(g_i^{-1}(\boldsymbol{\eta}_i)\Big)} \left|\frac{\partial}{\partial \boldsymbol{\eta}_i} g_i^{-1}(\boldsymbol{\eta}_i)\right|. \tag{2.7}$$

Proof. The Fisher information matrix of $\boldsymbol{\eta}$ is given by

$$\boldsymbol{H} = \boldsymbol{T}'\boldsymbol{\Sigma}\boldsymbol{T},$$

where $\boldsymbol{\Sigma}$ is the Fisher information matrix of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m)$ and

$$\boldsymbol{T} = diag\left(\frac{\partial}{\partial \boldsymbol{\eta}_1} g_1^{-1}(\boldsymbol{\eta}_1), \ldots, \frac{\partial}{\partial \boldsymbol{\eta}_m} g_m^{-1}(\boldsymbol{\eta}_m)\right).$$

The matrix $\boldsymbol{H}_{ii}^c$, which is derived by removing the rows and columns corresponding to $\boldsymbol{\eta}_i$ from $\boldsymbol{H}$, is of the form

$$\boldsymbol{H}_{ii}^c = \boldsymbol{T}_{ii}^{c'}\boldsymbol{\Sigma}_{ii}^c\boldsymbol{T}_{ii}^c,$$

where $\boldsymbol{\Sigma}_{ii}^c$ is the matrix which is derived by removing the rows and columns corresponding to $\boldsymbol{\theta}_i$ from $\boldsymbol{\Sigma}$, and

$$\boldsymbol{T}_{ii}^c = diag\left(\frac{\partial}{\partial \boldsymbol{\eta}_1} g_1^{-1}(\boldsymbol{\eta}_1), \ldots, \frac{\partial}{\partial \boldsymbol{\eta}_{i-1}} g_{i-1}^{-1}(\boldsymbol{\eta}_{i-1}), \frac{\partial}{\partial \boldsymbol{\eta}_{i+1}} g_{i+1}^{-1}(\boldsymbol{\eta}_{i+1}), \ldots, \frac{\partial}{\partial \boldsymbol{\eta}_m} g_m^{-1}(\boldsymbol{\eta}_m)\right).$$

Thus

$$\begin{aligned}
|\boldsymbol{H}| &= \prod_{k=1}^{m} \left|\frac{\partial}{\partial \boldsymbol{\eta}_k} g_k^{-1}(\boldsymbol{\eta}_k)\right|^2 |\boldsymbol{\Sigma}|, \\
|\boldsymbol{H}_{ii}^c| &= \prod_{j=1, j\neq i}^{m} \left|\frac{\partial}{\partial \boldsymbol{\eta}_j} g_j^{-1}(\boldsymbol{\eta}_j)\right|^2 |\boldsymbol{\Sigma}_{ii}^c|.
\end{aligned}$$

From Condition (2.3), it can be shown that

$$\begin{aligned}
\frac{|\boldsymbol{H}|}{|\boldsymbol{H}_{ii}^c|} &= \frac{\prod_{k=1}^{m} \left|\frac{\partial}{\partial \boldsymbol{\eta}_k} g_k^{-1}(\boldsymbol{\eta}_k)\right|^2}{\prod_{j=1, j\neq i}^{m} \left|\frac{\partial}{\partial \boldsymbol{\eta}_j} g_j^{-1}(\boldsymbol{\eta}_j)\right|^2} \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{ii}^c|} \\
&= \left|\frac{\partial}{\partial \boldsymbol{\eta}_i} g_i^{-1}(\boldsymbol{\eta}_i)\right|^2 f_{1i}\Big(g_i^{-1}(\boldsymbol{\eta}_i)\Big) f_{2i}\Big(g_i^{-1}(\boldsymbol{\eta}_i)^c\Big),
\end{aligned}$$

where

$$g_i^{-1}(\boldsymbol{\eta}_i)^c = \Big(g_1^{-1}(\boldsymbol{\eta}_1), \ldots, g_{i-1}^{-1}(\boldsymbol{\eta}_{i-1}), g_{i+1}^{-1}(\boldsymbol{\eta}_{i+1}), \ldots, g_m^{-1}(\boldsymbol{\eta}_m)\Big).$$

Thus we can write

$$\frac{|\boldsymbol{H}|}{|\boldsymbol{H}_{ii}^c|} = h_{1i}(\boldsymbol{\eta}_i)h_{2i}(\boldsymbol{\eta}_i^c),$$

where

$$\begin{aligned} h_{1i}(\boldsymbol{\eta}_i) &= f_{1i}\Big(g_i^{-1}(\boldsymbol{\eta}_i)\Big)\left|\frac{\partial}{\partial\boldsymbol{\eta}_i}g_i^{-1}(\boldsymbol{\eta}_i)\right|^2, \\ h_{2i}(\boldsymbol{\eta}_i^c) &= f_{2i}\Big(g_i^{-1}(\boldsymbol{\eta}_i)^c\Big). \end{aligned}$$

Hence, by Theorem 2.1, the independent reference prior for $\boldsymbol{\eta}$ is

$$\pi^R(\boldsymbol{\eta}) = \prod_{i=1}^m \pi_i^R(\boldsymbol{\eta}_i),$$

where the marginal reference priors for $\boldsymbol{\eta}_i$, $i = 1, \ldots, m$, are

$$\pi_i^R(\boldsymbol{\eta}_i) \propto \sqrt{h_{1i}(\boldsymbol{\eta}_i)} = \sqrt{f_{1i}\Big(g_i^{-1}(\boldsymbol{\eta}_i)\Big)}\left|\frac{\partial}{\partial\boldsymbol{\eta}_i}g_i^{-1}(\boldsymbol{\eta}_i)\right|.$$

The result then follows. $\square$

2.3 First Order Matching Priors

We propose a sufficient condition under which the independent reference prior, given by (2.4), is a first order matching prior. Thus we can easily prove that the independent reference prior is a first order matching prior without solving the differential equation given by (1.6).

Theorem 2.3 *Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$, where $\theta_i \in \boldsymbol{\Theta}_i \subset \mathbb{R}$. For fixed $i = 1, \cdots, m$, assume, for all $j = 1, \ldots, m$,*

$$\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{ij}^c|} = \begin{cases} f_{1i}(\theta_i)f_{2i}(\theta_i^c), & \textit{if } j = i, \\ \sqrt{f_{1j}(\theta_j)f_{2i}(\theta_i^c)}\, f_{3j}(\theta_j^c), & \textit{if } j \neq i, \end{cases} \tag{2.8}$$

where $\theta_l^c = (\theta_1, \ldots, \theta_{l-1}, \theta_{l+1}, \ldots, \theta_m)$, $l = i, j$, $\mathbf{\Sigma}$ is the Fisher information matrix of $\boldsymbol{\theta}$, $\mathbf{\Sigma}_{ij}^c$ is the matrix which is derived by removing the i^{th} row and j^{th} column from $\mathbf{\Sigma}$, and $f_{1j}f_{2i}$ for $j \neq i$ is a positive function of its argument. Then the independent reference prior $\pi^R(\boldsymbol{\theta})$ for $\boldsymbol{\theta}$, given by (2.4), is a first order matching prior for θ_i.

Proof. For fixed i, let $\psi = \psi(\boldsymbol{\theta}) = \theta_i$. By (2.8.3) of Datta and Mukerjee (2004), a first order matching prior $\pi^M = \pi^M(\theta_1, \ldots, \theta_m)$ for ψ satisfies the following differential equation,

$$\frac{\partial}{\partial\theta_1}(\xi_1\pi) + \cdots + \frac{\partial}{\partial\theta_m}(\xi_m\pi) = 0, \tag{2.9}$$

where

$$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)' = \frac{\mathbf{\Sigma}^{-1}\nabla\psi}{\sqrt{\nabla\psi'\mathbf{\Sigma}^{-1}\nabla\psi}}, \tag{2.10}$$

where $\nabla\psi = \left(\frac{\partial}{\partial\theta_1}\psi, \ldots, \frac{\partial}{\partial\theta_m}\psi\right)'$ and $\mathbf{\Sigma}$ is the Fisher information matrix of $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$. We need to show that the reference prior $\pi^R(\boldsymbol{\theta})$ for $\boldsymbol{\theta}$, given by (2.4), satisfies the equation (2.9). It is easy to see that $\nabla\psi = (0, \ldots, 0, 1, 0, \ldots, 0)'$, where 1 is the i-th element of $\nabla\psi$, and

$$\begin{aligned}
\boldsymbol{\xi} &= (\xi_1, \ldots, \xi_m)' \\
&= \sqrt{\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{ii}^c|}}\left\{(-1)^{i+1}\frac{|\mathbf{\Sigma}_{i1}^c|}{|\mathbf{\Sigma}|}, \ldots, \frac{|\mathbf{\Sigma}_{ii}^c|}{|\mathbf{\Sigma}|}, \ldots, (-1)^{i+m}\frac{|\mathbf{\Sigma}_{im}^c|}{|\mathbf{\Sigma}|}\right\}'.
\end{aligned}$$

From Condition (2.8), for $j = 1, \ldots, m$,

$$\begin{aligned}
\xi_j &= (-1)^{i+j}\sqrt{f_{1i}(\theta_i)f_{2i}(\theta_i^c)}\,\frac{|\mathbf{\Sigma}_{ij}^c|}{|\mathbf{\Sigma}|} \\
&= \begin{cases} \frac{1}{\sqrt{f_{1i}(\theta_i)f_{2i}(\theta_i^c)}}, & \text{if } j = i, \\ (-1)^{i+j}\sqrt{\frac{f_{1i}(\theta_i)}{f_{1j}(\theta_j)}}\,\frac{1}{f_{3j}(\theta_j^c)}, & \text{if } j \neq i. \end{cases}
\end{aligned}$$

Thus the differential equation (2.9) becomes

$$\frac{\partial}{\partial\theta_i}\left(\frac{\pi(\boldsymbol{\theta})}{\sqrt{f_{1i}(\theta_i)f_{2i}(\theta_i^c)}}\right) + \sum_{j=1, j\neq i}^{m}(-1)^{i+j}\frac{\partial}{\partial\theta_j}\left(\frac{\sqrt{f_{1i}(\theta_i)}\,\pi(\boldsymbol{\theta})}{\sqrt{f_{1j}(\theta_j)}\,f_{3j}(\theta_j^c)}\right) = 0.$$

Now, it can be shown that

$$\begin{aligned}
&\frac{\partial}{\partial\theta_i}\left(\frac{\pi^R(\boldsymbol{\theta})}{\sqrt{f_{1i}(\theta_i)f_{2i}(\theta_i^c)}}\right)+\sum_{j=1,j\neq i}^{m}(-1)^{i+j}\frac{\partial}{\partial\theta_j}\left(\frac{\sqrt{f_{1i}(\theta_i)}\,\pi^R(\boldsymbol{\theta})}{\sqrt{f_{1j}(\theta_j)}\,f_{3j}(\theta_j^c)}\right)\\
\propto\ &\frac{\partial}{\partial\theta_i}\left(\frac{\prod_{k=1}^m\sqrt{f_{1k}(\theta_k)}}{\sqrt{f_{1i}(\theta_i)f_{2i}(\theta_i^c)}}\right)+\sum_{j=1,j\neq i}^{m}(-1)^{i+j}\frac{\partial}{\partial\theta_j}\left(\frac{\sqrt{f_{1i}(\theta_i)}\,\prod_{k=1}^m\sqrt{f_{1k}(\theta_k)}}{\sqrt{f_{1j}(\theta_j)}\,f_{3j}(\theta_j^c)}\right)\\
=\ &\frac{\partial}{\partial\theta_i}\left(\frac{\prod_{k=1,k\neq i}^m\sqrt{f_{1k}(\theta_k)}}{\sqrt{f_{2i}(\theta_i^c)}}\right)+\sum_{j=1,j\neq i}^{m}(-1)^{i+j}\frac{\partial}{\partial\theta_j}\left(\frac{\sqrt{f_{1i}(\theta_i)}\,\prod_{k=1,k\neq j}^m\sqrt{f_{1k}(\theta_k)}}{f_{3j}(\theta_j^c)}\right)\\
=\ &0.
\end{aligned}$$

Hence the independent reference prior $\pi^R(\boldsymbol{\theta})$ for $\boldsymbol{\theta}$, given by (2.4), is a solution for the differential equation (2.9). The result then holds. $\square$

Corollary 2.2 *Suppose that in Condition (2.8), $|\boldsymbol{\Sigma}_{ij}^c| = 0$ for some $j \neq i$. The independent reference prior $\pi^R(\boldsymbol{\theta})$ for $\boldsymbol{\theta}$, given by (2.4), is a first order matching prior for θ_i.*

Proof. Clearly, if $|\boldsymbol{\Sigma}_{ij}^c| = 0$, then $\xi_j = 0$ for some $j \neq i$. Thus $\frac{\partial}{\partial\theta_j}(\xi_j\pi) = 0$ for any π. The result then follows. $\square$

Chapter 3

Examples

In this chapter, various examples of commonly used probability distributions are studied. We derive the independent reference priors by employing Theorem 2.1 and compare them with the reference priors. We also verify if the independent reference priors are also first order matching priors by applying Theorem 2.3. Consequently, the independent reference priors are shown to be the reference priors and first order matching priors in most of the examples. Note that most of the probability density functions, the Fisher information matrices and the reference priors in this chapter were provided by Yang and Berger (1997) unless other references are cited.

3.1 Binomial Model: Two Independent Samples

For fixed n_1 and n_2, let X_1 and X_2 be independent binomial random variables with the parameters (n_1, p_1) and (n_2, p_2), respectively. Then the joint density of (X_1, X_2) is

$$f(x_1, x_2|p_1, p_2) = \binom{n_1}{x_1} {p_1}^{x_1} (1 - p_1)^{n_1 - x_1} \binom{n_2}{x_2} {p_2}^{x_2} (1 - p_2)^{n_2 - x_2} \tag{3.1}$$

for $x_i \in \{0, 1, 2, \ldots, n_i\}$, $i = 1, 2$. Here $0 < p_1, p_2 < 1$. The Fisher information matrix of (p_1, p_2) is

$$\mathbf{\Sigma}(p_1, p_2) = \begin{pmatrix} \frac{n_1}{p_1(1-p_1)} & 0 \\ 0 & \frac{n_2}{p_2(1-p_2)} \end{pmatrix}. \tag{3.2}$$

Hence the marginal reference priors for p_1 and p_2 are

$$\pi_1^R(p_1) \propto \frac{1}{\sqrt{p_1(1-p_1)}},\ p_1 \in (0,1), \tag{3.3}$$

$$\pi_2^R(p_2) \propto \frac{1}{\sqrt{p_2(1-p_2)}},\ p_2 \in (0,1), \tag{3.4}$$

and the independent reference prior for (p_1, p_2) is

$$\pi^R(p_1, p_2) \propto \frac{1}{\sqrt{p_1(1-p_1)p_2(1-p_2)}}. \tag{3.5}$$

It is a first order matching prior for p_1 and p_2, and the reference prior for (p_1, p_2) when one of the parameters p_1 or p_2 is of interest and the other is nuisance parameter.

3.1.1 Two Binomial Proportions

Sun and Berger (1998) conducted objective Bayesian analysis by using the independent reference prior for the log-odds ratio of two binomial proportions in the example of a clinical trial: ECMO (extra corporeal membrane oxygenation). The ECMO example is described here: n_1 patients are given standard therapy and n_2 patients are treated with ECMO. The probability of success under standard therapy is p_1 and the probability of success under ECMO is p_2. Let X_1 be the number of survivors from standard therapy and X_2 be the number of survivors from ECMO. Then X_1 is a binomial random variable with parameters (n_1, p_1) and independently, X_2 is a binomial random variable with parameters (n_2, p_2). The main interest is to compare the two treatments. Then the log-odds ratio of p_1 and p_2, defined as $\delta = \eta_2 - \eta_1$ with $\eta_i = \log[p_i/(1-p_i)]$, $i = 1, 2$, is used for comparing them when η_1 is nuisance parameter. Under the assumption of the independence of δ and η_1, Sun and Berger (1998) obtained the marginal reference priors for δ ($\in \mathbb{R}$) and η_1 ($\in \mathbb{R}$), which are given by

$$\pi_1^R(\delta) \propto \exp\left(-\frac{1}{2\pi}\int_0^1 \{t(1-t)\}^{-0.5} \log\left[1 + \frac{n_1}{n_2}\{(1-t)e^{-\delta/2} + te^{\delta/2}\}^2\right] dt\right), \quad (3.6)$$

$$\pi_2^R(\eta_1) = \frac{e^{\eta_1/2}}{\pi(1+e^{\eta_1})}. \quad (3.7)$$

Consequently, the independent reference prior for (δ, η_1) is

$$\pi^R(\delta, \eta_1) \propto \frac{h(\delta)e^{\eta_1/2}}{1+e^{\eta_1}}, \quad (3.8)$$

where

$$h(\delta) = \exp\left(-\frac{1}{2\pi}\int_0^1 \{t(1-t)\}^{-0.5} \log\left[1 + \frac{n_1}{n_2}\{(1-t)e^{-\delta/2} + te^{\delta/2}\}^2\right] dt\right). \quad (3.9)$$

Now we compare the four priors for (δ, η_1) with respect to the frequentist matching property for δ and mean squared errors of the Bayes estimators of δ through simulation studies. The frequentist matching property is investigated by observing the frequentist coverage probabilities of the posterior credible interval for δ. The four priors considered here are constant prior, Jeffreys-rule prior, Cauchy prior and independent reference prior given by (3.8).

First, we compute the joint likelihood of (δ, η_1), which is given by

$$L^N(\delta, \eta_1) = \binom{n_1}{x_1}\left(\frac{e^{\eta_1}}{1+e^{\eta_1}}\right)^{x_1}\left(\frac{1}{1+e^{\eta_1}}\right)^{n_1-x_1}\binom{n_2}{x_2}\left(\frac{e^{\delta+\eta_1}}{1+e^{\delta+\eta_1}}\right)^{x_2}\left(\frac{1}{1+e^{\delta+\eta_1}}\right)^{n_2-x_2}$$

since the likelihood of (p_1, p_2) is given by (3.1) with $p_1 = \frac{e^{\eta_1}}{1+e^{\eta_1}}$ and $p_2 = \frac{e^{\delta+\eta_1}}{1+e^{\delta+\eta_1}}$.

We also derive the priors for (δ, η_1). The prior for (δ, η_1) corresponding to the constant prior for (p_1, p_2) is

$$\pi^C(\delta, \eta_1) \propto \frac{e^{\delta+2\eta_1}}{(1+e^{\eta_1})^2(1+e^{\delta+\eta_1})^2}, \quad (3.10)$$

the Jeffreys-rule prior for (δ, η_1) is

$$\pi^J(\delta, \eta_1) \propto \left[\frac{e^{\delta+2\eta_1}}{(1+e^{\eta_1})^2(1+e^{\delta+\eta_1})^2}\right]^{0.5}, \tag{3.11}$$

and the Cauchy prior for (δ, η_1) is

$$\pi^A(\delta, \eta_1) \propto \frac{1}{(1+\delta^2)(1+\eta_1^2)} \tag{3.12}$$

by assuming the independence of δ and η_1.

We then obtain the marginal posterior density functions for δ using the four priors. Let $n_S = x_1 + x_2$ and $n_F = n_1 + n_2 - n_S$. By using the transformations $\eta_1 = \log(\frac{t}{1-t})$ and $\delta = \log(\frac{u}{1-u})$, the marginal posterior density function for δ using the constant prior (3.10) is

$$\begin{aligned}
\pi^C(\delta|x_1, x_2) &= \frac{\int_{-\infty}^{\infty} L^N(\delta, \eta_1)\pi^C(\delta, \eta_1)d\eta_1}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} L^N(\delta, \eta_1)\pi^C(\delta, \eta_1)d\eta_1 d\delta} \\
&= \frac{(e^\delta)^{x_2+1}\int_0^1 t^{n_S+1}(1-t)^{n_F+1}\left(\frac{1}{1-t+e^\delta t}\right)^{n_2+2}dt}{\int_0^1\int_0^1 (\frac{u}{1-u})^{x_2+1}\frac{1}{u(1-u)}t^{n_S+1}(1-t)^{n_F+1}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2+2}dtdu},
\end{aligned}$$

the marginal posterior density for δ using the Jeffreys-rule prior (3.11) is

$$\pi^J(\delta|x_1, x_2) = \frac{(e^\delta)^{x_2+0.5}\int_0^1 t^{n_S}(1-t)^{n_F}\left(\frac{1}{1-t+e^\delta t}\right)^{n_2+1}dt}{\int_0^1\int_0^1 (\frac{u}{1-u})^{x_2+0.5}\frac{1}{u(1-u)}t^{n_S}(1-t)^{n_F}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2+1}dtdu},$$

the marginal posterior density function for δ using the Cauchy prior (3.12) is

$$\pi^A(\delta|x_1, x_2) = \frac{\frac{(e^\delta)^{x_2}}{1+\delta^2}\int_0^1 t^{n_S-1}(1-t)^{n_F-1}\left(\frac{1}{1-t+e^\delta t}\right)^{n_2}\frac{1}{1+\left\{\log(\frac{t}{1-t})\right\}^2}dt}{\int_0^1\int_0^1 \frac{(\frac{u}{1-u})^{x_2}}{1+\left\{\log(\frac{u}{1-u})\right\}^2}\frac{1}{u(1-u)}t^{n_S-1}(1-t)^{n_F-1}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2}\frac{1}{1+\left\{\log(\frac{t}{1-t})\right\}^2}dtdu},$$

and the marginal posterior density for δ using the independent reference prior (3.8) is

$$\pi^R(\delta|x_1, x_2) = \frac{(e^\delta)^{x_2}h(\delta)\int_0^1 t^{n_S-0.5}(1-t)^{n_F-0.5}\left(\frac{1}{1-t+e^\delta t}\right)^{n_2}dt}{\int_0^1 (\frac{u}{1-u})^{x_2}h(\log\{\frac{u}{1-u}\})\frac{1}{u(1-u)}\int_0^1 t^{n_S-0.5}(1-t)^{n_F-0.5}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2}dtdu},$$

where $h(\cdot)$ is given by (3.9).

Mean Squared Errors

Below are the analytical results for obtaining the mean squared error of the Bayes estimator of δ. Under the squared loss function, $L(\theta, a) = (\theta - a)^2$, the Bayes estimator of δ, $\hat{\delta}_l \equiv \hat{\delta}_l(x_1, x_2)$ is its posterior mean which is computed by $\int_{-\infty}^{\infty} \delta \pi^l(\delta|x_1, x_2) d\delta$, $l = C, J, A, R$. Thus by letting $\delta = \log\left(\frac{u}{1-u}\right)$,

$$\begin{aligned}
\widehat{\delta}_C &= \frac{\int_0^1 \int_0^1 \log\left(\frac{u}{1-u}\right)\left(\frac{u}{1-u}\right)^{x_2+1} \frac{1}{u(1-u)} t^{n_S+1}(1-t)^{n_F+1}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2+2} dtdu}{\int_0^1 \int_0^1 \left(\frac{u}{1-u}\right)^{x_2+1} \frac{1}{u(1-u)} t^{n_S+1}(1-t)^{n_F+1}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2+2} dtdu}, \\
\widehat{\delta}_J &= \frac{\int_0^1 \int_0^1 \log\left(\frac{u}{1-u}\right)\left(\frac{u}{1-u}\right)^{x_2+0.5} \frac{1}{u(1-u)} t^{n_S}(1-t)^{n_F}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2+1} dtdu}{\int_0^1 \int_0^1 \left(\frac{u}{1-u}\right)^{x_2+0.5} \frac{1}{u(1-u)} t^{n_S}(1-t)^{n_F}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2+1} dtdu}, \\
\widehat{\delta}_A &= \frac{\int_0^1 \int_0^1 \log\left(\frac{u}{1-u}\right)\frac{\left(\frac{u}{1-u}\right)^{x_2}}{1+\left\{\log\left(\frac{u}{1-u}\right)\right\}^2} \frac{1}{u(1-u)} t^{n_S-1}(1-t)^{n_F-1}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2} \frac{1}{1+\left\{\log\left(\frac{t}{1-t}\right)\right\}^2} dtdu}{\int_0^1 \int_0^1 \frac{\left(\frac{u}{1-u}\right)^{x_2}}{1+\left\{\log\left(\frac{u}{1-u}\right)\right\}^2} \frac{1}{u(1-u)} t^{n_S-1}(1-t)^{n_F-1}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2} \frac{1}{1+\left\{\log\left(\frac{t}{1-t}\right)\right\}^2} dtdu}, \\
\widehat{\delta}_R &= \frac{\int_0^1 \log\left(\frac{u}{1-u}\right)\left(\frac{u}{1-u}\right)^{x_2} h\left(\log\left\{\frac{u}{1-u}\right\}\right)\frac{1}{u(1-u)} \int_0^1 t^{n_S-0.5}(1-t)^{n_F-0.5}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2} dtdu}{\int_0^1 \left(\frac{u}{1-u}\right)^{x_2} h\left(\log\left\{\frac{u}{1-u}\right\}\right)\frac{1}{u(1-u)} \int_0^1 t^{n_S-0.5}(1-t)^{n_F-0.5}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2} dtdu},
\end{aligned}$$

where $h(\cdot)$ is given by (3.9).

Hence, the mean squared error is given by

$$MSE_l = E\left[\left(\widehat{\delta}_l - \delta\right)^2\right] = \sum_{x_1=0}^{n_1} \sum_{x_2=0}^{n_2} \left(\widehat{\delta}_l - \delta\right)^2 L^N(\delta, \eta_1) \tag{3.13}$$

for $l = C, J, A, R$.

Computing (3.13) is straightforward if n_1 and n_2 are small. However the following approximation is proposed for large n_1 and n_2. For fixed (δ^*, η_1^*), $p_1^* = \frac{e^{\eta_1^*}}{1+e^{\eta_1^*}}$ and $p_2^* = \frac{e^{\delta^*+\eta_1^*}}{1+e^{\delta^*+\eta_1^*}}$ are obtained. Then, for fixed n_1 and n_2, two sets of independent binomial random variables, $x_1^{(k)}|p_1^*$ and $x_2^{(k)}|p_2^*$ with (n_1, p_1^*) and (n_2, p_2^*), respectively, are generated for $k = 1, \ldots, K$. For the simulated $x_1^{(k)}$ and $x_2^{(k)}$, $\widehat{\delta}_l^{(k)}$, $l = C, J, A, R$, can be calculated by using the above equations. Then $\left(\widehat{\delta}_l^{(k)} - \delta^*\right)^2$, $k = 1, \ldots, K$, are computed for $l = C, J, A, R$. Thus the estimate of MSE_l is

$$\widehat{MSE}_l = \frac{1}{K}\sum_{k=1}^{K}\left(\widehat{\delta}_l^{(k)} - \delta^*\right)^2$$

for $l = C, J, A, R$.

The results are shown in Table 3.1 in the end of Section 3.1. A prior which has small mean squared errors is desirable. We considered small ($n_1 = n_2 = 10$) and large sample sizes ($n_1 = n_2 = 50$). We then chose $\delta^* = -2, -1, 0, 1, 2$ and $\eta_1^* = -2, -1, 0, 1, 2$ for each δ^*, and ran $K = 5000$ replicates for each set of (δ^*, η_1^*). It is observed that the mean squared errors obtained by using the Jeffreys-rule prior and the independent reference prior are larger than those using the constant prior and the Cauchy prior for both small and large samples. However the differences are much smaller for large samples. Thus the constant prior and the Cauchy prior might perform better than the Jeffreys-rule prior and the independent reference prior in the inference of δ with respect to the mean squared errors.

Frequentist Coverage Probabilities

We explain how to compute the frequentist coverage probability of the one-sided posterior credible interval for δ. For any $\alpha \in (0, 1)$, let $q_\alpha^l(x_1, x_2)$ be the posterior α-quantile of δ, i.e. $P(\delta \le q_\alpha^l(x_1, x_2)|x_1, x_2) = \alpha$ for $l = C, J, A, R$. Then the frequentist coverage probability of the one-sided $(\alpha \times 100)\%$ posterior credible interval $(-\infty, q_\alpha^l(x_1, x_2))$ is defined as, for $l = C, J, A, R$,

$$P_{(\delta,\eta_1)}(\delta \le q_\alpha^l(x_1, x_2)) = \sum_{x_1=0}^{n_1}\sum_{x_2=0}^{n_2} I\left\{\delta \le q_\alpha^l(x_1, x_2)\right\} L^N(\delta, \eta_1),$$

where $I\{\cdot\}$ is the indicator function. It is desired that the frequentist coverage probability is close to α. It could be difficult to compute the frequentist coverage probability if $q_\alpha^l(x_1, x_2)$,

$l = C, J, A, R$, is found first. Alternatively, we first consider, for fixed (δ^*, η_1^*),

$$\Big\{(x_1, x_2) : \delta^* \le q_\alpha^l(x_1, x_2)\Big\} = \left\{(x_1, x_2) : \int_{-\infty}^{\delta^*} \pi^l(\delta|x_1, x_2)d\delta < \alpha\right\}$$

for $l = C, J, A, R$. Then the frequentist coverage probability can be approximated as follows.

For the simulated $x_1^{(k)}$ and $x_2^{(k)}$, $k = 1, \ldots, K$, which are generated as the previous section on mean squared errors, the posterior density function, $\pi^l(\delta|x_1^{(k)}, x_2^{(k)})$, $l = C, J, A, R$, can be computed. Then the estimate of $P_{(\delta,\eta_1)}(\delta \le q_\alpha^l(x_1, x_2))$ is given by

$$\widehat{P}_{(\delta,\eta_1)}(\delta^* \le q_\alpha^l(x_1, x_2)) = \frac{1}{K}\sum_{k=1}^{K} I\left\{\int_{-\infty}^{\delta^*} \pi^l(\delta|x_1^{(k)}, x_2^{(k)})d\delta < \alpha\right\}$$

for $l = C, J, A, R$. It is shown that by letting $\delta = \log\left(\frac{u}{1-u}\right)$,

$$\begin{aligned}
&\int_{-\infty}^{\delta^*} \pi^C(\delta|x_1, x_2)d\delta \\
&= \frac{\int_0^{\frac{e^{\delta^*}}{1+e^{\delta^*}}}\int_0^1 \left(\frac{u}{1-u}\right)^{x_2+1}\frac{1}{u(1-u)}t^{n_S+1}(1-t)^{n_F+1}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2+2}dtdu}{\int_0^1\int_0^1 \left(\frac{u}{1-u}\right)^{x_2+1}\frac{1}{u(1-u)}t^{n_S+1}(1-t)^{n_F+1}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2+2}dtdu}, \\
&\int_{-\infty}^{\delta^*} \pi^J(\delta|x_1, x_2)d\delta \\
&= \frac{\int_0^{\frac{e^{\delta^*}}{1+e^{\delta^*}}}\int_0^1 \left(\frac{u}{1-u}\right)^{x_2+0.5}\frac{1}{u(1-u)}t^{n_S}(1-t)^{n_F}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2+1}dtdu}{\int_0^1\int_0^1 \left(\frac{u}{1-u}\right)^{x_2+0.5}\frac{1}{u(1-u)}t^{n_S}(1-t)^{n_F}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2+1}dtdu}, \\
&\int_{-\infty}^{\delta^*} \pi^A(\delta|x_1, x_2)d\delta \\
&= \frac{\int_0^{\frac{e^{\delta^*}}{1+e^{\delta^*}}}\int_0^1 \frac{\left(\frac{u}{1-u}\right)^{x_2}}{1+\left\{\log\left(\frac{u}{1-u}\right)\right\}^2}\frac{1}{u(1-u)}t^{n_S-1}(1-t)^{n_F-1}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2}\frac{1}{1+\left\{\log\left(\frac{t}{1-t}\right)\right\}^2}dtdu}{\int_0^1\int_0^1 \frac{\left(\frac{u}{1-u}\right)^{x_2}}{1+\left\{\log\left(\frac{u}{1-u}\right)\right\}^2}\frac{1}{u(1-u)}t^{n_S-1}(1-t)^{n_F-1}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2}\frac{1}{1+\left\{\log\left(\frac{t}{1-t}\right)\right\}^2}dtdu}, \\
&\int_{-\infty}^{\delta^*} \pi^R(\delta|x_1, x_2)d\delta \\
&= \frac{\int_0^{\frac{e^{\delta^*}}{1+e^{\delta^*}}} \left(\frac{u}{1-u}\right)^{x_2}h(\log\{\frac{u}{1-u}\})\frac{1}{u(1-u)}\int_0^1 t^{n_S-0.5}(1-t)^{n_F-0.5}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2}dtdu}{\int_0^1 \left(\frac{u}{1-u}\right)^{x_2}h(\log\{\frac{u}{1-u}\})\frac{1}{u(1-u)}\int_0^1 t^{n_S-0.5}(1-t)^{n_F-0.5}\left(\frac{1}{1-t+\frac{u}{1-u}t}\right)^{n_2}dtdu},
\end{aligned}$$

where $h(\cdot)$ is given by (3.9).

The output is given in Table 3.2–3.4 in the end of Section 3.1. Table 3.2–3.4 displays the frequentist coverage probabilities of the one-sided $(\alpha \times 100)\%$ posterior credible interval for δ when $\alpha = 0.05, 0.5, 0.95$, respectively. Recall that we want a prior whose frequentist coverage probabilities are close to α. We considered small $(n_1 = n_2 = 10)$ and large sample sizes $(n_1 = n_2 = 50)$. We then chose $\delta^* = -2, -1, 0, 1, 2$ and $\eta_1^* = -2, -1, 0, 1, 2$ for each δ^*, and ran $K = 5000$ replicates for each set of (δ^*, η_1^*). From Table 3.2–3.4, it is roughly seen that the frequentist coverage probabilities computed by using the Jeffreys-rule prior and the independent reference prior are much closer to α than those using the constant prior and the Cauchy prior for both small and large samples. It is also observed that the frequentist coverage probabilities derived by using the constant prior are closer to α than those using the Cauchy prior. It is clear that the frequentist coverage probabilities are consistently much closer to α for large samples than small samples. Hence the Jeffreys-rule prior and the independent reference prior could be better priors for the inference of δ than the constant prior which is better than the Cauchy prior with respect to the frequentist matching property. This conclusion is the opposite of the one obtained when considering the mean squared errors.

Table 3.1: Mean Squared Errors

		$n_1 = n_2 = 10$				$n_1 = n_2 = 50$			
δ^*	η_1^*	π^C	π^J	π^A	π^R	π^C	π^J	π^A	π^R
	-2	2.0207	2.2071	1.1014	1.0707	0.7234	1.6563	1.7609	1.5684
	-1	1.1600	1.9900	1.6708	1.5447	0.5274	1.0233	1.2081	0.9465
-2	0	0.9538	2.0119	2.1184	1.5439	0.2906	0.3562	0.3591	0.3645
	1	1.0413	1.8522	1.6140	1.2355	0.2029	0.2215	0.2372	0.2131
	2	0.9688	1.9893	1.5156	1.0943	0.2761	0.3478	0.3553	0.2877
	-2	1.1580	2.7084	1.3875	2.2053	0.6331	1.1753	0.9977	1.1996
	-1	1.0099	2.2334	1.6232	1.9648	0.2935	0.3504	0.2895	0.3674
-1	0	0.9081	1.3929	0.9198	1.2212	0.1744	0.1852	0.1603	0.1836
	1	0.8969	1.4533	0.7190	1.0455	0.1892	0.2010	0.1962	0.1939
	2	1.0845	2.3403	1.2581	1.6868	0.3181	0.3771	0.3337	0.3424
	-2	1.1739	2.9737	1.5390	2.6305	0.3989	0.5046	0.2951	0.4931
	-1	0.9795	1.7037	0.8091	1.4414	0.2009	0.2150	0.1427	0.2113
0	0	0.8008	1.0526	0.4174	0.9482	0.1575	0.1641	0.1058	0.1609
	1	1.0224	1.7590	0.8079	1.4546	0.1944	0.2080	0.1384	0.2042
	2	1.2169	3.0536	1.5991	2.6568	0.4260	0.5493	0.3124	0.5335
	-2	1.0892	2.3760	1.1688	1.6387	0.3050	0.3614	0.3160	0.3256
	-1	0.8925	1.4068	0.7051	1.0327	0.1888	0.2004	0.1967	0.1934
1	0	0.8907	1.4054	0.9467	1.2359	0.1809	0.1913	0.1659	0.1893
	1	1.0986	2.4028	1.6893	2.0667	0.3060	0.3775	0.3235	0.3905
	2	1.1179	2.5853	1.3270	2.1399	0.6171	1.1596	0.9922	1.1855
	-2	0.9753	2.0445	1.4677	1.1112	0.2883	0.3738	0.3666	0.2958
	-1	0.9909	1.7259	1.5285	1.1936	0.2005	0.2161	0.2415	0.2090
2	0	0.9761	2.0967	2.2125	1.5867	0.2838	0.3438	0.3473	0.3523
	1	1.1905	2.0721	1.7404	1.5976	0.5000	0.9685	1.1343	0.9104
	2	2.0456	2.3697	1.1405	1.2183	0.6822	1.5668	1.6855	1.4915

Table 3.2: Frequentist Coverage Probabilities of One-sided 5% Posterior Credible Interval for δ

		$n_1 = n_2 = 10$				$n_1 = n_2 = 50$			
δ^*	η_1^*	π^C	π^J	π^A	π^R	π^C	π^J	π^A	π^R
	-2	0.0440	0.0440	0.0445	0.0610	0.1000	0.0690	0.1005	0.0415
	-1	0.0530	0.0520	0.0625	0.0440	0.0695	0.0530	0.0800	0.0515
-2	0	0.0540	0.0440	0.1125	0.0440	0.0515	0.0495	0.0865	0.0480
	1	0.0600	0.0510	0.1950	0.0485	0.0460	0.0360	0.0910	0.0370
	2	0.0615	0.0435	0.2980	0.0595	0.0430	0.0390	0.1240	0.0415
	-2	0.0275	0.0275	0.0275	0.0265	0.0560	0.0505	0.0560	0.0475
	-1	0.0300	0.0300	0.0340	0.0315	0.0520	0.0435	0.0570	0.0460
-1	0	0.0425	0.0425	0.0920	0.0340	0.0525	0.0495	0.0760	0.0435
	1	0.0295	0.0295	0.1505	0.0400	0.0535	0.0505	0.1225	0.0530
	2	0.0385	0.0385	0.2140	0.0965	0.0685	0.0620	0.1460	0.0680
	-2	0.0290	0.0290	0.0005	0.0055	0.0450	0.0510	0.0140	0.0375
	-1	0.0425	0.0425	0.0055	0.0325	0.0520	0.0535	0.0230	0.0440
0	0	0.0550	0.0550	0.0180	0.0580	0.0435	0.0435	0.0335	0.0445
	1	0.0505	0.0505	0.0420	0.0810	0.0435	0.0465	0.0440	0.0580
	2	0.0370	0.0370	0.0370	0.1090	0.0490	0.0540	0.0490	0.0665
	-2	0.0095	0.0320	0.0005	0.0085	0.0425	0.0610	0.0090	0.0420
	-1	0.0270	0.0765	0.0065	0.0245	0.0500	0.0630	0.0140	0.0460
1	0	0.0300	0.0655	0.0215	0.0585	0.0415	0.0535	0.0210	0.0530
	1	0.0090	0.0315	0.0085	0.0625	0.0365	0.0490	0.0340	0.0580
	2	0.0000	0.0015	0.0000	0.0120	0.0240	0.0570	0.0415	0.0685
	-2	0.0035	0.0440	0.0000	0.0175	0.0370	0.0585	0.0195	0.0430
	-1	0.0165	0.0810	0.0085	0.0795	0.0395	0.0440	0.0190	0.0555
2	0	0.0015	0.0590	0.0015	0.0565	0.0375	0.0610	0.0300	0.0570
	1	0.0000	0.0025	0.0000	0.0030	0.0150	0.0685	0.0520	0.0730
	2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0050	0.0020	0.0120

Table 3.3: Frequentist Coverage Probabilities of One-sided 50% Posterior Credible Interval for δ

		$n_1 = n_2 = 10$				$n_1 = n_2 = 50$			
δ^*	η_1^*	π^C	π^J	π^A	π^R	π^C	π^J	π^A	π^R
	-2	0.9150	0.7175	0.9150	0.4055	0.6675	0.4935	0.5700	0.4295
	-1	0.6755	0.4880	0.6820	0.3730	0.5920	0.4885	0.5550	0.4360
-2	0	0.5770	0.4355	0.6630	0.4180	0.5520	0.4810	0.6010	0.4810
	1	0.5945	0.4235	0.7730	0.4460	0.5130	0.4875	0.6750	0.4995
	2	0.5855	0.4450	0.8825	0.5595	0.5510	0.4900	0.7135	0.5380
	-2	0.7475	0.5275	0.5465	0.3545	0.5565	0.5085	0.6255	0.4485
	-1	0.5475	0.5020	0.6235	0.4445	0.5190	0.4725	0.5345	0.4770
-1	0	0.5245	0.5245	0.7190	0.4695	0.5095	0.4800	0.6310	0.4840
	1	0.4905	0.4890	0.8345	0.5410	0.5175	0.4935	0.7000	0.5085
	2	0.5565	0.5110	0.9190	0.6445	0.5520	0.5080	0.7690	0.5350
	-2	0.3540	0.3605	0.3540	0.3675	0.5230	0.4590	0.4225	0.4290
	-1	0.4390	0.4875	0.4195	0.4150	0.5085	0.5060	0.4685	0.4705
0	0	0.4685	0.5460	0.5410	0.4635	0.5035	0.5040	0.5030	0.5200
	1	0.5590	0.5225	0.5885	0.5830	0.4985	0.4920	0.5405	0.5585
	2	0.6510	0.6450	0.6520	0.6370	0.4660	0.5275	0.5740	0.5535
	-2	0.4650	0.5045	0.0930	0.3400	0.4520	0.5070	0.2420	0.4675
	-1	0.5035	0.5040	0.1690	0.4685	0.4685	0.5015	0.2935	0.5025
1	0	0.4900	0.4910	0.2860	0.5290	0.4775	0.5020	0.3735	0.5180
	1	0.4600	0.5040	0.3840	0.5365	0.4645	0.5205	0.4505	0.5265
	2	0.2430	0.4740	0.4525	0.6495	0.4575	0.5110	0.3830	0.5275
	-2	0.4305	0.5735	0.1250	0.4465	0.4715	0.5325	0.3085	0.4705
	-1	0.3965	0.5620	0.2260	0.5630	0.4710	0.4955	0.3190	0.4980
2	0	0.4295	0.5630	0.3395	0.5610	0.4425	0.5075	0.3955	0.5310
	1	0.3495	0.5225	0.3440	0.6210	0.4200	0.5245	0.4465	0.5410
	2	0.0875	0.3030	0.0875	0.6065	0.3285	0.4980	0.4165	0.5290

Table 3.4: Frequentist Coverage Probabilities of One-sided 95% Posterior Credible Interval for δ

		$n_1 = n_2 = 10$				$n_1 = n_2 = 50$			
δ^*	η_1^*	π^C	π^J	π^A	π^R	π^C	π^J	π^A	π^R
	-2	1.0000	1.0000	1.0000	1.0000	1.0000	0.9950	0.9970	0.9885
	-1	1.0000	0.9965	1.0000	0.9965	0.9815	0.9290	0.9440	0.9195
-2	0	0.9985	0.9485	0.9985	0.9400	0.9585	0.9345	0.9620	0.9345
	1	0.9735	0.9075	0.9870	0.9275	0.9485	0.9450	0.9740	0.9480
	2	0.9965	0.9520	0.9990	0.9815	0.9640	0.9430	0.9815	0.9610
	-2	1.0000	0.9980	1.0000	0.9865	0.9735	0.9370	0.9585	0.9240
	-1	0.9930	0.9740	0.9930	0.9315	0.9610	0.9390	0.9635	0.9415
-1	0	0.9695	0.9310	0.9815	0.9265	0.9615	0.9485	0.9770	0.9550
	1	0.9645	0.9175	0.9940	0.9750	0.9490	0.9410	0.9865	0.9495
	2	0.9915	0.9710	0.9995	0.9920	0.9620	0.9460	0.9885	0.9615
	-2	0.9660	0.9660	0.9665	0.9010	0.9485	0.9425	0.9485	0.9260
	-1	0.9560	0.9560	0.9600	0.9150	0.9585	0.9505	0.9545	0.9395
0	0	0.9395	0.9395	0.9830	0.9475	0.9645	0.9645	0.9730	0.9555
	1	0.9490	0.9490	0.9940	0.9585	0.9570	0.9540	0.9815	0.9520
	2	0.9630	0.9630	0.9980	0.9945	0.9440	0.9420	0.9820	0.9615
	-2	0.9625	0.9625	0.7900	0.8980	0.9430	0.9485	0.8525	0.9390
	-1	0.9645	0.9645	0.8385	0.9620	0.9420	0.9450	0.8760	0.9500
1	0	0.9630	0.9630	0.9215	0.9590	0.9445	0.9465	0.9160	0.9515
	1	0.9610	0.9610	0.9540	0.9565	0.9435	0.9525	0.9365	0.9590
	2	0.9775	0.9775	0.9775	0.9745	0.9500	0.9530	0.9500	0.9640
	-2	0.9400	0.9565	0.7190	0.9305	0.9550	0.9570	0.8740	0.9410
	-1	0.9390	0.9500	0.7985	0.9435	0.9455	0.9580	0.9015	0.9625
2	0	0.9485	0.9645	0.8920	0.9585	0.9445	0.9480	0.9080	0.9515
	1	0.9470	0.9475	0.9245	0.9455	0.9420	0.9580	0.9250	0.9510
	2	0.9405	0.9405	0.9405	0.9365	0.9065	0.9425	0.9065	0.9595

3.2 Bivariate Binomial Model

Crowder and Sweeting (1989) considered the following bivariate binomial density

$$f(x_1, x_2|p_1, p_2) = \binom{n}{x_1} {p_1}^{x_1}(1-p_1)^{n-x_1}\binom{x_1}{x_2}{p_2}^{x_2}(1-p_2)^{x_1-x_2},$$

where for fixed n, $x_1 \in \{0, 1, 2, \ldots, n\}, x_2 \in \{0, 1, 2, \ldots, x_1\}$. Here $0 < p_1, p_2 < 1$. Then the Fisher information matrix of (p_1, p_2) is

$$\mathbf{\Sigma}(p_1, p_2) = n \begin{pmatrix} \frac{1}{p_1(1-p_1)} & 0 \\ 0 & \frac{p_1}{p_2(1-p_2)} \end{pmatrix}. \tag{3.14}$$

Hence the marginal reference priors for p_1 and p_2 are

$$\pi_1^R(p_1) \;\propto\; \frac{1}{\sqrt{p_1(1-p_1)}},\;\; p_1 \in (0,1), \tag{3.15}$$

$$\pi_2^R(p_2) \;\propto\; \frac{1}{\sqrt{p_2(1-p_2)}},\;\; p_2 \in (0,1), \tag{3.16}$$

and the independent reference prior for (p_1, p_2) is

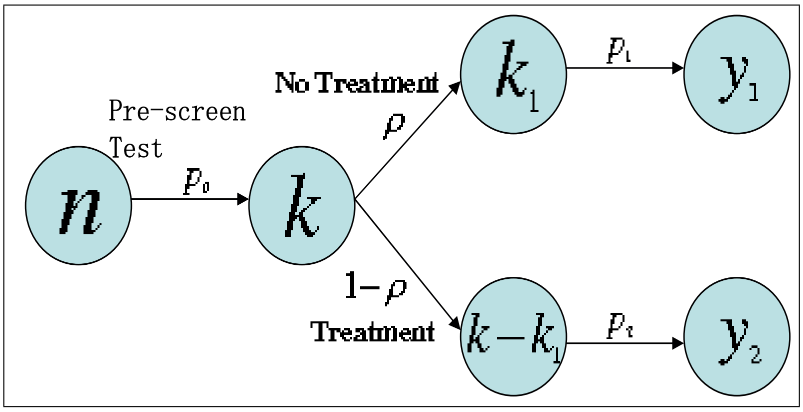
$$\pi^R(p_1, p_2) \propto \frac{1}{\sqrt{p_1(1-p_1)p_2(1-p_2)}}. \tag{3.17}$$

It is also a first order matching prior for p_1 and p_2, and the reference prior for (p_1, p_2) when one of the parameters p_1 or p_2 is the parameter of interest and the other is nuisance parameter.

3.3 Two Binomial Proportions with Pre-screen Test

Extending the ECMO example of Sun and Berger (1998), we now mainly derive two independent reference priors for the log-odds ratio of two binomial proportions when an initial screen test is taken in a clinical trial. We consider the following two cases.

Figure 3.1: Diagram for Case I



3.3.1 Case I

Suppose that n individuals are chosen for a clinical trial. A pre-screen test is applied with probability of survival p_0. Then the individuals who are alive after the screen test are randomly given no treatment with probability ρ and a treatment with probability 1-ρ. Let p_1 be the probability of success for the non-treatment (control) group and p_2 be the probability of success for the treatment group. Figure 3.1 shows the diagram for this example. Then the probability density function is given by

$$f(k, k_1, y_1, y_2 \mid p_0, p_1, p_2, \rho) = \binom{n}{k} {p_0}^k (1-p_0)^{n-k} \binom{k}{k_1} \rho^{k_1} (1-\rho)^{k-k_1} \\ \times \binom{k_1}{y_1} {p_1}^{y_1} (1-p_1)^{k_1-y_1} \binom{k-k_1}{y_2} {p_2}^{y_2} (1-p_2)^{k-k_1-y_2}, \tag{3.18}$$

where for fixed n, $k \in \{0,1,2,\ldots,n\}, k_1 \in \{0,1,2,\ldots,k\}, y_1 \in \{0,1,2,\ldots,k_1\}, y_2 \in \{0,1,2,\ldots,k-k_1\}$. Thus the Fisher information matrix of (p_0, p_1, p_2, ρ) is

$$\mathbf{\Sigma}(p_0, p_1, p_2, \rho) = n \begin{pmatrix} \frac{1}{p_0(1-p_0)} & 0 & 0 & 0 \\ 0 & \frac{p_0\rho}{p_1(1-p_1)} & 0 & 0 \\ 0 & 0 & \frac{p_0(1-\rho)}{p_2(1-p_2)} & 0 \\ 0 & 0 & 0 & \frac{p_0}{\rho(1-\rho)} \end{pmatrix}. \tag{3.19}$$

The following result is easy and the proof is omitted.

Proposition 3.1 *Consider the model (3.18).*

(a) The marginal reference priors for p_0, p_1, p_2 and ρ are

$$\pi_1^R(p_0) \propto \frac{1}{\sqrt{p_0(1-p_0)}}, \; p_0 \in (0,1), \tag{3.20}$$

$$\pi_2^R(p_1) \propto \frac{1}{\sqrt{p_1(1-p_1)}}, \; p_1 \in (0,1), \tag{3.21}$$

$$\pi_3^R(p_2) \propto \frac{1}{\sqrt{p_2(1-p_2)}}, \; p_2 \in (0,1), \tag{3.22}$$

$$\pi_4^R(\rho) \propto \frac{1}{\sqrt{\rho(1-\rho)}}, \; \rho \in (0,1). \tag{3.23}$$

(b) The independent reference prior for (p_0, p_1, p_2, ρ) is

$$\pi^R(p_0, p_1, p_2, \rho) \propto \frac{1}{\sqrt{p_0(1-p_0)p_1(1-p_1)p_2(1-p_2)\rho(1-\rho)}}. \tag{3.24}$$

(c) The prior in (b) is a first order matching prior for p_0, p_1, p_2 and ρ.

(d) The prior in (b) is the one-at-a-time reference prior for (p_0, p_1, p_2, ρ) with any ordering.

Now consider the log-odds ratio of p_1 and p_2 defined as

$$\delta = \eta_2 - \eta_1, \tag{3.25}$$

where

$$\eta_i = \log\left(\frac{p_i}{1-p_i}\right), \quad i = 1, 2. \tag{3.26}$$

It is the interest parameter to compare the treatment and control group. Then the Fisher information matrix of $(\delta, \eta_1, p_0, \rho)$ is

$$\mathbf{\Sigma}(\delta, \eta_1, p_0, \rho) = \begin{pmatrix} B & B & 0 & 0 \\ B & B+C & 0 & 0 \\ 0 & 0 & \frac{n}{p_0(1-p_0)} & 0 \\ 0 & 0 & 0 & \frac{np_0}{\rho(1-\rho)} \end{pmatrix}, \tag{3.27}$$

where

$$B = \frac{np_0(1-\rho)e^{\delta+\eta_1}}{(1+e^{\delta+\eta_1})^2} \quad \text{and} \quad C = \frac{np_0\rho e^{\eta_1}}{(1+e^{\eta_1})^2}.$$

Thus

$$\begin{aligned}
|\mathbf{\Sigma}| &= BC\,\frac{n^2}{(1-p_0)(1-\rho)\rho}, \\
|\mathbf{\Sigma}_{11}^c| &= (B+C)\,\frac{n^2}{(1-p_0)(1-\rho)\rho}, \quad |\mathbf{\Sigma}_{22}^c| = B\,\frac{n^2}{(1-p_0)(1-\rho)\rho}, \\
|\mathbf{\Sigma}_{33}^c| &= BC\,\frac{np_0}{\rho(1-\rho)}, \quad |\mathbf{\Sigma}_{44}^c| = BC\,\frac{n}{p_0(1-p_0)}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{BC}{B+C} = \frac{np_0\rho(1-\rho)e^{\delta+\eta_1}}{\rho(1+e^{\delta+\eta_1})^2 + e^{\delta}(1-\rho)(1+e^{\eta_1})^2}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} &= C = \frac{np_0\rho e^{\eta_1}}{(1+e^{\eta_1})^2}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} &= \frac{n}{p_0(1-p_0)}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{44}^c|} &= \frac{np_0}{\rho(1-\rho)}.
\end{aligned}$$

We note that $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{ii}^c|$, $i = 2,3,4$, satisfy Condition (2.3) but $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{11}^c|$ does not. Thus we cannot apply Theorem 2.1 to this problem. We use the iterative algorithm directly to derive the independent reference prior for $(\delta, \eta_1, p_0, \rho)$.

Proposition 3.2 *Consider the model (3.18) with the new parameterization (3.25) and (3.26). Then the marginal reference priors for δ, η_1, p_0 and ρ are*

$$\begin{aligned}
\pi_1^R(\delta) &\propto \exp\left\{\frac{1}{2}\int_0^1\int_{-\infty}^{\infty} \pi_2^R(\eta_1)\pi_4^R(\rho)\log[h_1(\delta,\eta_1,\rho)]d\eta_1 d\rho\right\}, \ \delta \in \mathbb{R}, && (3.28) \\
\pi_2^R(\eta_1) &= \frac{e^{\eta_1/2}}{\pi(1+e^{\eta_1})}, \ \eta_1 \in \mathbb{R}, && (3.29) \\
\pi_3^R(p_0) &= \frac{1}{\pi\sqrt{p_0(1-p_0)}}, \ p_0 \in (0,1), && (3.30) \\
\pi_4^R(\rho) &= \frac{1}{\pi\sqrt{\rho(1-\rho)}}, \ \rho \in (0,1), && (3.31)
\end{aligned}$$

where

$$h_1(\delta, \eta_1, \rho) = \frac{\rho(1-\rho)e^{\delta+\eta_1}}{\rho(1+e^{\delta+\eta_1})^2 + e^{\delta}(1-\rho)(1+e^{\eta_1})^2}.$$

Consequently, the independent reference prior for $(\delta, \eta_1, p_0, \rho)$ *is*

$$\pi^R(\delta, \eta_1, p_0, \rho) = \pi_1^R(\delta)\pi_2^R(\eta_1)\pi_3^R(p_0)\pi_4^R(\rho). \tag{3.32}$$

Proof. Because $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{ii}^c|$, $i = 2, 3, 4$, satisfy Condition (2.3), (3.29)–(3.31) hold immediately. It is easily shown that $\pi_2^R(\eta_1)$, $\pi_3^R(p_0)$ and $\pi_4^R(\rho)$ are proper. Thus we need to apply *Algorithm B* to derive $\pi_1^R(\delta)$ since $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{11}^c|$ does not meet Condition (2.3). Then

$$\begin{aligned}\pi_1^R(\delta) &\propto \exp\left\{\frac{1}{2}\int_0^1\int_0^1\int_{-\infty}^{\infty}\pi_2^R(\eta_1)\pi_3^R(p_0)\pi_4^R(\rho)\log\left(\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|}\right)d\eta_1 dp_0 d\rho\right\}\\ &= \exp\left\{\frac{1}{2}\int_0^1\int_0^1\int_{-\infty}^{\infty}\pi_2^R(\eta_1)\pi_3^R(p_0)\pi_4^R(\rho)\log[h_1^*(\delta, \eta_1, p_0, \rho)]d\eta_1 dp_0 d\rho\right\},\end{aligned}$$

where

$$h_1^*(\delta, \eta_1, p_0, \rho) = \frac{p_0\rho(1-\rho)e^{\delta+\eta_1}}{\rho(1+e^{\delta+\eta_1})^2 + e^{\delta}(1-\rho)(1+e^{\eta_1})^2}.$$

Clearly,

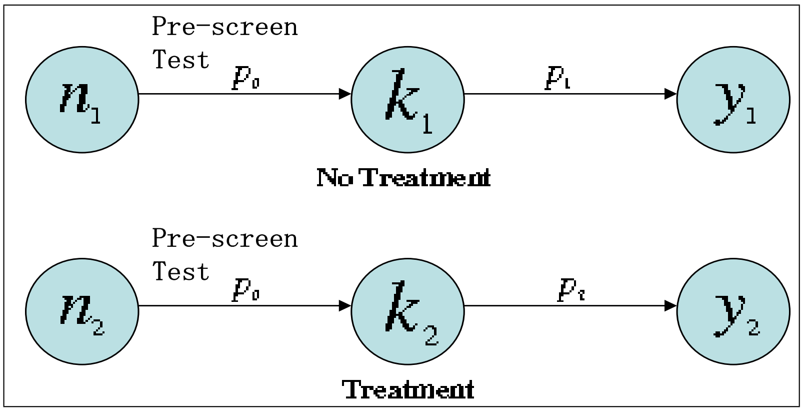
$$\pi_1^R(\delta) \propto \exp\left\{\frac{1}{2}\int_0^1 \pi_3^R(p_0)\log(p_0)dp_0 + \frac{1}{2}\int_0^1\int_{-\infty}^{\infty}\pi_2^R(\eta_1)\pi_4^R(\rho)\log[h_1(\delta, \eta_1, \rho)]d\eta_1 d\rho\right\}.$$

The result holds. □

3.3.2 Case II

Suppose that two groups of n_1 and n_2 individuals are selected for a clinical trial. First, an initial screen test is conducted to the group of n_1 individuals with probability of survival p_0. Then the individuals who are alive after the screen test are given no treatment. Next,

Figure 3.2: Diagram for Case II



the same initial screen test is conducted to the group of n_2 individuals with probability of survival p_0. Then the individuals who are alive after the screen test are given a treatment. Let p_1 be the probability of success under no treatment and p_2 be the probability of success under treatment. The diagram for this case is given in Figure 3.2. Then the probability density function is given by

$$f(k_1, y_1, k_2, y_2 \mid p_0, p_1, p_2) = \binom{n_1}{k_1} {p_0}^{k_1}(1-p_0)^{n_1-k_1} \binom{k_1}{y_1} {p_1}^{y_1}(1-p_1)^{k_1-y_1} \\ \times \binom{n_2}{k_2} {p_0}^{k_2}(1-p_0)^{n_2-k_2} \binom{k_2}{y_2} {p_2}^{y_2}(1-p_2)^{k_2-y_2}, \quad (3.33)$$

where for fixed n_1 and n_2, $k_1 \in \{0, 1, 2, \ldots, n_1\}, y_1 \in \{0, 1, 2, \ldots, k_1\}, k_2 \in \{0, 1, 2, \ldots, n_2\}, y_2 \in \{0, 1, 2, \ldots, k_2\}$. Thus the Fisher information matrix of (p_0, p_1, p_2) is

$$\mathbf{\Sigma}(p_0, p_1, p_2) = \begin{pmatrix} \frac{n_1+n_2}{p_0(1-p_0)} & 0 & 0 \\ 0 & \frac{n_1 p_0}{p_1(1-p_1)} & 0 \\ 0 & 0 & \frac{n_2 p_0}{p_2(1-p_2)} \end{pmatrix}. \quad (3.34)$$

The following proposition is easy and the proof is omitted.

Proposition 3.3 *Consider the model (3.33).*

(a) The marginal reference priors for p_0, p_1 and p_2 are

$$\pi_1^R(p_0) \propto \frac{1}{\sqrt{p_0(1-p_0)}}, \; p_0 \in (0,1), \tag{3.35}$$

$$\pi_2^R(p_1) \propto \frac{1}{\sqrt{p_1(1-p_1)}}, \; p_1 \in (0,1), \tag{3.36}$$

$$\pi_3^R(p_2) \propto \frac{1}{\sqrt{p_2(1-p_2)}}, \; p_2 \in (0,1). \tag{3.37}$$

(b) The independent reference prior for (p_0, p_1, p_2) is

$$\pi^R(p_0, p_1, p_2) \propto \frac{1}{\sqrt{p_0(1-p_0)p_1(1-p_1)p_2(1-p_2)}}. \tag{3.38}$$

(c) The prior in (b) is a first order matching prior for p_0, p_1 and p_2.

(d) The prior in (b) is the one-at-a-time reference prior for (p_0, p_1, p_2) with any ordering.

As stated in the previous case, we are interested in $\delta = \eta_2 - \eta_1$ given by (3.25) and (3.26). The Fisher information matrix of (δ, η_1, p_0) is

$$\mathbf{\Sigma}(\delta, \eta_1, p_0) = \begin{pmatrix} B & B & 0 \\ B & B+C & 0 \\ 0 & 0 & \frac{n_1+n_2}{p_0(1-p_0)} \end{pmatrix}, \tag{3.39}$$

where

$$B = \frac{n_2 p_0 e^{\delta+\eta_1}}{(1+e^{\delta+\eta_1})^2} \quad \text{and} \quad C = \frac{n_1 p_0 e^{\eta_1}}{(1+e^{\eta_1})^2}.$$

Clearly,

$$|\mathbf{\Sigma}| = BC\frac{n_1+n_2}{p_0(1-p_0)}, \quad |\mathbf{\Sigma}_{11}^c| = (B+C)\frac{n_1+n_2}{p_0(1-p_0)},$$
$$|\mathbf{\Sigma}_{22}^c| = B\frac{n_1+n_2}{p_0(1-p_0)}, \quad |\mathbf{\Sigma}_{33}^c| = BC.$$

Then

$$\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} = \frac{BC}{B+C} = \frac{n_1 n_2 p_0 e^{\delta+\eta_1}}{n_1(1+e^{\delta+\eta_1})^2 + n_2 e^{\delta}(1+e^{\eta_1})^2},$$

$$\begin{aligned}
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{22}^c|} &= C = \frac{n_1 p_0 e^{\eta_1}}{(1+e^{\eta_1})^2}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{33}^c|} &= \frac{n_1+n_2}{p_0(1-p_0)}.
\end{aligned}$$

It is similar to *Case I* that $|\boldsymbol{\Sigma}|/|\boldsymbol{\Sigma}_{ii}^c|$, $i=2,3$, satisfy Condition (2.3) but $|\boldsymbol{\Sigma}|/|\boldsymbol{\Sigma}_{11}^c|$ does not. Thus we cannot apply Theorem 2.1 to this problem either. We use the iterative algorithm to find the independent reference prior for (δ, η_1, p_0).

Proposition 3.4 *Consider the model (3.33) with the new parameterization (3.25) and (3.26). Then the marginal reference priors for δ, η_1 and p_0 are*

$$\pi_1^R(\delta) \propto \exp\left\{\frac{1}{2}\int_{-\infty}^{\infty} \pi_2^R(\eta_1)\log[h_2(\delta,\eta_1)]d\eta_1\right\},\ \delta \in \mathbb{R}, \tag{3.40}$$

$$\pi_2^R(\eta_1) = \frac{e^{\eta_1/2}}{\pi(1+e^{\eta_1})},\ \eta_1 \in \mathbb{R}, \tag{3.41}$$

$$\pi_3^R(p_0) = \frac{1}{\pi\sqrt{p_0(1-p_0)}},\ p_0 \in (0,1), \tag{3.42}$$

where

$$h_2(\delta,\eta_1) = \frac{e^{\delta+\eta_1}}{n_1(1+e^{\delta+\eta_1})^2 + n_2 e^{\delta}(1+e^{\eta_1})^2}.$$

Consequently, the independent reference prior for (δ, η_1, p_0) is

$$\pi^R(\delta,\eta_1,p_0) = \pi_1^R(\delta)\pi_2^R(\eta_1)\pi_3^R(p_0). \tag{3.43}$$

Proof. It is clear that $|\boldsymbol{\Sigma}|/|\boldsymbol{\Sigma}_{22}^c|$ and $|\boldsymbol{\Sigma}|/|\boldsymbol{\Sigma}_{33}^c|$ satisfy Condition (2.3) so that (3.41) and (3.42) hold immediately. It is easy to see that $\pi_2^R(\eta_1)$ and $\pi_3^R(p_0)$ are proper. Thus we need to apply *Algorithm B* to derive $\pi_1^R(\delta)$ since $|\boldsymbol{\Sigma}|/|\boldsymbol{\Sigma}_{11}^c|$ does not meet Condition (2.3). Then

$$\begin{aligned}
\pi_1^R(\delta) &\propto \exp\left\{\frac{1}{2}\int_0^1\int_{-\infty}^{\infty} \pi_2^R(\eta_1)\pi_3^R(p_0)\log\left(\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}^c|}\right)d\eta_1 dp_0\right\} \\
&= \exp\left\{\frac{1}{2}\int_0^1\int_{-\infty}^{\infty} \pi_2^R(\eta_1)\pi_3^R(p_0)\log[h_2^*(\delta,\eta_1,p_0)]d\eta_1 dp_0\right\},
\end{aligned}$$

where

$$h_2^*(\delta, \eta_1, p_0) = \frac{p_0 e^{\delta+\eta_1}}{n_1(1+e^{\delta+\eta_1})^2 + n_2 e^{\delta}(1+e^{\eta_1})^2}.$$

Clearly,

$$\pi_1^R(\delta) \propto \exp\left\{\frac{1}{2}\int_0^1 \pi_3^R(p_0)\log(p_0)dp_0 + \frac{1}{2}\int_{-\infty}^{\infty} \pi_2^R(\eta_1)\log[h_1(\delta,\eta_1)]d\eta_1\right\}.$$

The result then holds. □

3.4 Exponential Model: Two Independent Samples

Let X_1 and X_2 be independent exponential random variables with means $1/\theta_1$ and $1/\theta_2$, respectively. Here $\theta_i > 0$ are unknown. The joint density of (X_1, X_2) is

$$f(x_1, x_2 \mid \theta_1, \theta_2) = \theta_1 \exp(-x_1\theta_1)\theta_2 \exp(-x_2\theta_2), \quad x_1, x_2 \geq 0.$$

It is easy to compute the Fisher information matrix of (θ_1, θ_2), which is given by

$$\mathbf{\Sigma}(\theta_1, \theta_2) = \begin{pmatrix} \frac{1}{\theta_1^2} & 0 \\ 0 & \frac{1}{\theta_2^2} \end{pmatrix}. \tag{3.44}$$

Hence the marginal reference priors for θ_1 and θ_2 are

$$\pi_1^R(\theta_1) \propto \frac{1}{\theta_1}, \ \theta_1 > 0, \tag{3.45}$$

$$\pi_2^R(\theta_2) \propto \frac{1}{\theta_2}, \ \theta_2 > 0, \tag{3.46}$$

and the independent reference prior for (θ_1, θ_2) is

$$\pi^R(\theta_1, \theta_2) \propto \frac{1}{\theta_1\theta_2}. \tag{3.47}$$

It is a first order matching prior for θ_1 and θ_2, and also the reference prior for (θ_1, θ_2) when one of the parameters θ_1 or θ_2 is of interest and the other is nuisance parameter.

Alternatively, let $\phi = \frac{\theta_2}{\theta_1+\theta_2}$, the proportion explained by the mean of X_1 in total mean of X_1 and X_2, and $\omega = \theta_1 + \theta_2$. The Fisher information matrix of (ϕ, ω) is

$$\mathbf{\Sigma}(\phi, \omega) = \begin{pmatrix} \frac{\phi^2+(1-\phi)^2}{\phi^2(1-\phi)^2} & \frac{1-2\phi}{\omega\phi(1-\phi)} \\ \frac{1-2\phi}{\omega\phi(1-\phi)} & \frac{2}{\omega^2} \end{pmatrix}. \tag{3.48}$$

Clearly,

$$\begin{aligned} |\mathbf{\Sigma}| &= \frac{1}{\omega^2\phi^2(1-\phi)^2}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{2}{\omega^2}, \\ |\mathbf{\Sigma}_{22}^c| &= \frac{\phi^2+(1-\phi)^2}{\phi^2(1-\phi)^2}, \quad |\mathbf{\Sigma}_{12}^c| = |\mathbf{\Sigma}_{21}^c| = \frac{1-2\phi}{\omega\phi(1-\phi)}. \end{aligned}$$

Then

$$\begin{aligned} \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{1}{2\phi^2(1-\phi)^2}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} &= \frac{1}{\omega^2}\left\{\frac{1}{\phi^2+(1-\phi)^2}\right\}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = \frac{1}{\omega\phi(1-\phi)(1-2\phi)}. \end{aligned}$$

Hence the marginal reference priors for ϕ and ω are

$$\pi_1^R(\phi) \propto \frac{1}{\phi(1-\phi)}, \ \phi \in (0,1), \tag{3.49}$$

$$\pi_2^R(\omega) \propto \frac{1}{\omega}, \ \omega > 0, \tag{3.50}$$

and the independent reference prior for (ϕ, ω) is

$$\pi^R(\phi, \omega) \propto \frac{1}{\omega\phi(1-\phi)}. \tag{3.51}$$

It is a first order matching prior for ϕ, and the reference prior for (ϕ, ω) when ϕ is the parameter of interest and ω is nuisance parameter.

We could consider the third set of parameters. Let $\phi = \frac{\theta_1}{\theta_2}$, the ratio of two means, and $\omega = \theta_1\theta_2$. Then the Fisher information matrix of (ϕ, ω) is

$$\boldsymbol{\Sigma}(\phi, \omega) = \frac{1}{4}\begin{pmatrix} \frac{2}{\phi^2} & 0 \\ 0 & \frac{2}{\omega^2} \end{pmatrix}. \tag{3.52}$$

Hence the marginal reference priors for ϕ and ω are

$$\pi_1^R(\phi) \quad \propto \quad \frac{1}{\phi}, \ \phi > 0, \tag{3.53}$$

$$\pi_2^R(\omega) \quad \propto \quad \frac{1}{\omega}, \ \omega > 0, \tag{3.54}$$

and the independent reference prior for (ϕ, ω) is

$$\pi^R(\phi, \omega) \propto \frac{1}{\phi\omega}. \tag{3.55}$$

It is a first order matching prior for ϕ and ω. By Datta and M. Ghosh (1995), the independent reference prior for (ϕ, ω) is the same as the reference prior for (ϕ, ω) when one of the parameters ϕ or ω is the interest and the other is nuisance parameter.

3.5 Gamma Model

Consider the gamma density

$$f(x \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x), \quad x > 0.$$

Here $\alpha > 0$ and $\beta > 0$ are unknown paramaters. The Fisher information matrix of (α, β) is

$$\boldsymbol{\Sigma}(\alpha, \beta) = \begin{pmatrix} \xi(\alpha) & -\frac{1}{\beta} \\ -\frac{1}{\beta} & \frac{\alpha}{\beta^2} \end{pmatrix}, \tag{3.56}$$

where $\xi(\alpha) = \sum_{i=0}^{\infty} (x+i)^{-2}$. It is easy to see that

$$|\boldsymbol{\Sigma}| = \frac{\xi(\alpha)\alpha - 1}{\beta^2}, \quad |\boldsymbol{\Sigma}_{11}^c| = \frac{\alpha}{\beta^2}, \quad |\boldsymbol{\Sigma}_{22}^c| = \xi(\alpha), \quad |\boldsymbol{\Sigma}_{12}^c| = |\boldsymbol{\Sigma}_{21}^c| = -\frac{1}{\beta}.$$

Then we have

$$\begin{aligned}
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}^c|} &= \frac{\xi(\alpha)\alpha - 1}{\alpha}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{22}^c|} &= \frac{1}{\beta^2}\left\{\frac{\xi(\alpha)\alpha - 1}{\xi(\alpha)}\right\}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{12}^c|} &= \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{21}^c|} = -\frac{\xi(\alpha)\alpha - 1}{\beta}.
\end{aligned}$$

Hence the marginal reference priors for α and β are

$$\pi_1^R(\alpha) \propto \sqrt{\frac{\xi(\alpha)\alpha - 1}{\alpha}},\ \alpha > 0, \tag{3.57}$$

$$\pi_2^R(\beta) \propto \frac{1}{\beta},\ \beta > 0. \tag{3.58}$$

The independent reference prior for (α, β) is

$$\pi^R(\alpha, \beta) \propto \frac{\sqrt{\xi(\alpha)\alpha - 1}}{\sqrt{\alpha}\beta}. \tag{3.59}$$

It is a first order matching prior for α, and also the reference prior for (α, β) when α is the parameter of interest and β is nuisance parameter.

We can also consider alternative reparameterization (α, μ) for the gamma model, where $\mu = E(x \mid \alpha, \mu)$. The density is

$$f(x \mid \alpha, \mu) = \frac{\alpha^\alpha}{\mu^\alpha \Gamma(\alpha)} x^{\alpha - 1} \exp\left(-\frac{\alpha}{\mu} x\right).$$

Then the Fisher information matrix of (α, μ) is

$$\boldsymbol{\Sigma}(\alpha, \mu) = \begin{pmatrix} \frac{\xi(\alpha)\alpha - 1}{\alpha} & 0 \\ 0 & \frac{\alpha}{\mu^2} \end{pmatrix}. \tag{3.60}$$

Hence the marginal reference priors for α and μ are

$$\pi_1^R(\alpha) \propto \sqrt{\frac{\xi(\alpha)\alpha - 1}{\alpha}},\ \alpha > 0, \tag{3.61}$$

$$\pi_2^R(\mu) \propto \frac{1}{\mu},\ \mu > 0, \tag{3.62}$$

and the independent reference prior for (α, μ) is

$$\pi^R(\alpha, \mu) \propto \frac{\sqrt{\xi(\alpha)\alpha - 1}}{\sqrt{\alpha}\mu}. \tag{3.63}$$

It is a first order matching prior for α and μ, and the reference prior for (α, μ) when one of α or μ is the parameter of interest and the other is nuisance parameter.

3.6 Inverse Gaussian Model

For $x > 0, \alpha > 0, \psi > 0$, the inverse Gaussian density is

$$f(x|\alpha, \psi) = \left(\frac{\alpha}{2\pi}\right)^{1/2} \frac{1}{x^{3/2}} \exp\left\{-\frac{\alpha x}{2}\left(\psi - \frac{1}{x}\right)^2\right\}.$$

Then the Fisher information matrix of (α, ψ) is

$$\mathbf{\Sigma}(\alpha, \psi) = \begin{pmatrix} \frac{1}{2\alpha^2} & 0 \\ 0 & \frac{\alpha}{\psi} \end{pmatrix}. \tag{3.64}$$

Hence the marginal reference priors for α and ψ are

$$\pi_1^R(\alpha) \propto \frac{1}{\alpha}, \ \alpha > 0, \tag{3.65}$$

$$\pi_2^R(\psi) \propto \frac{1}{\sqrt{\psi}}, \ \psi > 0, \tag{3.66}$$

and the independent reference prior for (α, ψ) is

$$\pi^R(\alpha, \psi) \propto \frac{1}{\alpha\sqrt{\psi}}. \tag{3.67}$$

It is a first order matching prior for α and ψ, and the reference prior for (α, ψ) when one of the parameters α or ψ is of interest and the other is nuisance parameter.

Now we consider alternative parameterization, $\theta = \sqrt{\frac{\alpha}{\psi}}$ and $\beta = \sqrt{\alpha\psi}$. The Fisher information matrix of (θ, β) is

$$\mathbf{\Sigma}(\theta, \beta) = \frac{1}{2}\begin{pmatrix} \frac{1+2\beta^2}{\theta^2} & \frac{1-2\beta^2}{\theta\beta} \\ \frac{1-2\beta^2}{\theta\beta} & \frac{1+2\beta^2}{\beta^2} \end{pmatrix}. \tag{3.68}$$

Thus

$$|\mathbf{\Sigma}| = \frac{2}{\theta^2}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{1+2\beta^2}{\beta^2}, \quad |\mathbf{\Sigma}_{22}^c| = \frac{1+2\beta^2}{\theta^2}, \quad |\mathbf{\Sigma}_{12}^c| = |\mathbf{\Sigma}_{21}^c| = \frac{1-2\beta^2}{2\theta\beta},$$

and then

$$\begin{aligned} \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{1}{\theta^2}\left(\frac{2\beta^2}{1+2\beta^2}\right), \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} &= \frac{2}{1+2\beta^2}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = \frac{4\beta}{\theta(1-2\beta^2)}. \end{aligned}$$

Hence the marginal reference priors for θ and β are

$$\pi_1^R(\theta) \propto \frac{1}{\theta}, \ \theta > 0, \tag{3.69}$$

$$\pi_2^R(\beta) \propto \frac{1}{\sqrt{1+2\beta^2}}, \ \beta > 0, \tag{3.70}$$

and the independent reference prior for (θ, β) is

$$\pi^R(\theta, \beta) \propto \frac{1}{\theta\sqrt{1+2\beta^2}}. \tag{3.71}$$

It is also a first order matching prior for β, and the reference prior for (θ, β) when β is the interest parameter and θ is nuisance parameter.

We now consider the third parameterization for the inverse Gaussian density. Then for $x > 0, \mu > 0, \sigma > 0$, the inverse Gaussian density is rewritten as

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{x^{3/2}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2\mu^2 x}\right\}.$$

The Fisher information matrix of (μ, σ^2) is

$$\mathbf{\Sigma}(\mu, \sigma^2) = \begin{pmatrix} \frac{1}{\mu^3\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{pmatrix}. \tag{3.72}$$

Hence the marginal reference priors for μ and σ^2 are

$$\pi_1^R(\mu) \propto \frac{1}{\mu^{3/2}}, \ \mu > 0, \tag{3.73}$$

$$\pi_2^R(\sigma^2) \propto \frac{1}{\sigma^2}, \ \sigma^2 > 0, \tag{3.74}$$

and the independent reference prior for (μ, σ^2) is

$$\pi^R(\mu, \sigma^2) \propto \frac{1}{\mu^{3/2}\sigma^2}. \tag{3.75}$$

It is a first order matching prior for μ and σ^2. By Datta and M. Ghosh (1995), the independent reference prior for (μ, σ^2) is equivalent to the reference prior for (μ, σ^2) when one of the parameters μ or σ^2 is of interest and the other is nuisance parameter.

3.7 Lognormal Model

The lognormal density is, for $x > 0$,

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{x} \exp\left\{ -\frac{(\log x - \mu)^2}{2\sigma^2} \right\},$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$ are unknown parameters. Then the Fisher information matrix of (μ, σ) is

$$\mathbf{\Sigma}(\mu, \sigma) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix}. \tag{3.76}$$

Hence the marginal reference priors for μ and σ are

$$\pi_1^R(\mu) \propto 1, \ \mu \in \mathbb{R}, \tag{3.77}$$

$$\pi_2^R(\sigma) \propto \frac{1}{\sigma}, \ \sigma > 0, \tag{3.78}$$

and the independent reference prior for (μ, σ) is

$$\pi^R(\mu, \sigma) \propto \frac{1}{\sigma}. \tag{3.79}$$

It is a first order matching prior for μ and σ, and the reference prior for (μ, σ) when one of μ or σ is the parameter of interest and the other is nuisance parameter.

Now we consider alternative parameterization. Let $\theta = \exp\left(\mu + \frac{\sigma^2}{2}\right)$, the mean, and $\beta = \sigma^2$. Then the Fisher information matrix of (θ, β) is

$$\mathbf{\Sigma}(\theta, \beta) = \frac{1}{2}\begin{pmatrix} \frac{1}{\theta^2\beta} & -\frac{1}{2\theta\beta} \\ -\frac{1}{2\theta\beta} & \frac{\beta+2}{4\beta^2} \end{pmatrix}. \tag{3.80}$$

It is easy to compute

$$|\mathbf{\Sigma}| = \frac{1}{2\theta^2\beta^3}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{\beta+2}{4\beta^2}, \quad |\mathbf{\Sigma}_{22}^c| = \frac{1}{\theta^2\beta^2}, \quad |\mathbf{\Sigma}_{12}^c| = |\mathbf{\Sigma}_{21}^c| = -\frac{1}{4\theta\beta}.$$

Then

$$\begin{aligned} \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{1}{\theta^2}\left\{\frac{2}{\beta(\beta+2)}\right\}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} &= \frac{1}{2\beta^2}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = -\frac{2}{\theta\beta^2}. \end{aligned}$$

Hence the marginal reference priors for θ and β are

$$\pi_1^R(\theta) \propto \frac{1}{\theta}, \ \theta > 0, \tag{3.81}$$

$$\pi_2^R(\beta) \propto \frac{1}{\beta}, \ \beta > 0, \tag{3.82}$$

and the independent reference prior for (θ, β) is

$$\pi^R(\theta, \beta) \propto \frac{1}{\theta\beta}. \tag{3.83}$$

It is a first order matching prior for β. Furthermore, it is also the reference prior when β is the parameter of interest and θ is the nuisance parameter.

3.8 Normal Model

For $x \in \mathbb{R}$, the normal density is

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}.$$

Here $\mu \in \mathbb{R}$ is a unknown mean and $\sigma^2 > 0$ is a unknown variance. Then the Fisher information matrix of (μ, σ) is

$$\mathbf{\Sigma}(\mu, \sigma) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix}. \tag{3.84}$$

Hence the marginal reference priors for μ and σ are

$$\pi_1^R(\mu) \propto 1, \ \mu \in \mathbb{R}, \tag{3.85}$$

$$\pi_2^R(\sigma) \propto \frac{1}{\sigma}, \ \sigma > 0, \tag{3.86}$$

and the independent reference prior for (μ, σ) is

$$\pi^R(\mu, \sigma) \propto \frac{1}{\sigma}. \tag{3.87}$$

It is also a first order matching prior for μ and σ, and the reference prior for (μ, σ) when one of the parameters μ or σ is the parameter of interest and the other is nuisance parameter.

3.9 Normal Model: Two Independent Samples

3.9.1 Unequal Variances

Let X_1 and X_2 be independent normal random variables with means μ_1 and μ_2, and variances σ_1^2 and σ_2^2. Then for $x_i \in \mathbb{R}, \mu_i \in \mathbb{R}, \sigma_i > 0$, $i = 1, 2$, the joint density is

$$f(x_1, x_2|\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}\right\} \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}\right\}.$$

Then the Fisher information matrix of $(\mu_1, \mu_2, \sigma_1, \sigma_2)$ is

$$\mathbf{\Sigma}(\mu_1, \mu_2, \sigma_1, \sigma_2) = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 & 0 & 0 \\ 0 & \frac{1}{\sigma_2^2} & 0 & 0 \\ 0 & 0 & \frac{2}{\sigma_1^2} & 0 \\ 0 & 0 & 0 & \frac{2}{\sigma_2^2} \end{pmatrix}. \tag{3.88}$$

Hence the marginal reference priors for μ_1, μ_2, σ_1 and σ_2 are

$$\pi_1^R(\mu_1) \propto 1, \ \mu_1 \in \mathbb{R}, \tag{3.89}$$

$$\pi_2^R(\mu_2) \propto 1, \ \mu_2 \in \mathbb{R}, \tag{3.90}$$

$$\pi_3^R(\sigma_1) \propto \frac{1}{\sigma_1}, \ \sigma_1 > 0, \tag{3.91}$$

$$\pi_4^R(\sigma_2) \propto \frac{1}{\sigma_2}, \ \sigma_2 > 0, \tag{3.92}$$

and the independent reference prior for $(\mu_1, \mu_2, \sigma_1, \sigma_2)$ is

$$\pi^R(\mu_1, \mu_2, \sigma_1, \sigma_2) \propto \frac{1}{\sigma_1\sigma_2}. \tag{3.93}$$

It is a first order matching prior for μ_1, μ_2, σ_1 and σ_2, and also the one-at-a-time reference prior for $(\mu_1, \mu_2, \sigma_1, \sigma_2)$ with any ordering.

Alternatively, let $\theta = \mu_1\mu_2$, the product of two means, and $\omega = \sqrt{\frac{\mu_2}{\mu_1}}$ when $\mu_1, \mu_2 > 0$. Then the Fisher information matrix of $(\theta, \omega, \sigma_1, \sigma_2)$ is

$$\mathbf{\Sigma}(\theta, \omega, \sigma_1, \sigma_2) = \begin{pmatrix} \frac{\omega^4\sigma_1^2+\sigma_2^2}{4\theta\omega^2\sigma_1^2\sigma_2^2} & \frac{\omega^4\sigma_1^2-\sigma_2^2}{2\omega^3\sigma_1^2\sigma_2^2} & 0 & 0 \\ \frac{\omega^4\sigma_1^2-\sigma_2^2}{2\omega^3\sigma_1^2\sigma_2^2} & \frac{\theta(\omega^4\sigma_1^2+\sigma_2^2)}{\omega^4\sigma_1^2\sigma_2^2} & 0 & 0 \\ 0 & 0 & \frac{2}{\sigma_1^2} & 0 \\ 0 & 0 & 0 & \frac{2}{\sigma_2^2} \end{pmatrix}. \tag{3.94}$$

Thus

$$\begin{aligned} |\mathbf{\Sigma}| &= \frac{4}{\omega^2\sigma_1^4\sigma_2^4}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{4\theta(\omega^4\sigma_1^2+\sigma_2^2)}{\omega^4\sigma_1^4\sigma_2^4}, \quad |\mathbf{\Sigma}_{22}^c| = \frac{\omega^4\sigma_1^2+\sigma_2^2}{\theta\omega^2\sigma_1^4\sigma_2^4}, \\ |\mathbf{\Sigma}_{33}^c| &= \frac{2}{\omega^2\sigma_1^2\sigma_2^4}, \quad |\mathbf{\Sigma}_{44}^c| = \frac{2}{\omega^2\sigma_1^4\sigma_2^2}, \end{aligned}$$

and then

$$\begin{aligned} \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{\omega^2}{\theta(\omega^4\sigma_1^2+\sigma_2^2)}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} &= \frac{4\theta}{\omega^4\sigma_1^2+\sigma_2^2}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} &= \frac{2}{\sigma_1^2}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{44}^c|} = \frac{2}{\sigma_2^2}. \end{aligned}$$

It is clear that $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{ii}^c|$, $i = 1, 3, 4$, satisfy Condition (2.3) but $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{22}^c|$ does not. Thus we cannot apply Theorem 2.1 to this example. We use the iterative algorithm to compute the independent reference prior for $(\theta, \omega, \sigma_1, \sigma_2)$.

Proposition 3.5 *Let* $\{[1/\sqrt{j}, \sqrt{j}],\ j = 1, 2, \ldots\}$ *be an increasing sequence of compact subsets of* $(0, \infty)$ *for* σ_1 *and* $\{[1/\sqrt{2j}, \sqrt{2j}],\ j = 1, 2, \ldots\}$ *for* σ_2*. Then the marginal reference priors for* θ*,* ω*,* σ_1 *and* σ_2 *are*

$$\pi_1^R(\theta) \propto \frac{1}{\sqrt{\theta}}, \ \theta > 0, \tag{3.95}$$

$$\pi_2^R(\omega) \propto \frac{1}{\omega}, \ \omega > 0, \tag{3.96}$$

$$\pi_3^R(\sigma_1) \propto \frac{1}{\sigma_1}, \ \sigma_1 > 0, \tag{3.97}$$
$$\pi_4^R(\sigma_2) \propto \frac{1}{\sigma_2}, \ \sigma_2 > 0. \tag{3.98}$$

Consequently, the independent reference prior for $(\theta, \omega, \sigma_1, \sigma_2)$ *is*

$$\pi^R(\theta, \omega, \sigma_1, \sigma_2) \propto \frac{1}{\sqrt{\theta}\omega\sigma_1\sigma_2}. \tag{3.99}$$

Proof. Clearly, (3.95), (3.97) and (3.98) hold since $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{ii}^c|$, $i = 1, 3, 4$, satisfy Condition (2.3). It is easily shown that $\pi_1^R(\theta)$, $\pi_3^R(\sigma_1)$ and $\pi_4^R(\sigma_2)$ are improper. Thus we need an argument of compact sets and use *Algorithm B′* to derive $\pi_2^R(\omega)$ since $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{22}^c|$ does not meet Condition (2.3). Choose $\{[a_j, b_j],\ j = 1, 2, \ldots\}$ as an increasing sequence of compact subsets of $(0, \infty)$ for θ, where $a_j \to 0$ and $b_j \to \infty$ as $j \to \infty$. Also choose $\{[1/\sqrt{j}, \sqrt{j}],\ j = 1, 2, \ldots\}$ as an increasing sequence of compact subsets of $(0, \infty)$ for σ_1 and $\{[1/\sqrt{2j}, \sqrt{2j}],\ j = 1, 2, \ldots\}$ for σ_2. Then by applying *Algorithm B′*, it can be seen that

$$\begin{aligned}
\pi_{2j}(\omega) &\propto \exp\left\{\frac{\int_{1/\sqrt{2j}}^{\sqrt{2j}}\int_{1/\sqrt{j}}^{\sqrt{j}}\int_{a_j}^{b_j}\pi_1^R(\theta)\pi_3^R(\sigma_1)\pi_4^R(\sigma_2)\log\left(\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|}\right)d\theta d\sigma_1 d\sigma_2}{2\int_{1/\sqrt{2j}}^{\sqrt{2j}}\int_{1/\sqrt{j}}^{\sqrt{j}}\int_{a_j}^{b_j}\pi_1^R(\theta)\pi_3^R(\sigma_1)\pi_4^R(\sigma_2)d\theta d\sigma_1 d\sigma_2}\right\} \\
&= \exp\left\{\frac{\int_{1/\sqrt{2j}}^{\sqrt{2j}}\int_{1/\sqrt{j}}^{\sqrt{j}}\int_{a_j}^{b_j}\frac{1}{\sqrt{\theta}\sigma_1\sigma_2}\log\left(\frac{4\theta}{\omega^4\sigma_1^2+\sigma_2^2}\right)d\theta d\sigma_1 d\sigma_2}{2\int_{1/\sqrt{2j}}^{\sqrt{2j}}\int_{1/\sqrt{j}}^{\sqrt{j}}\int_{a_j}^{b_j}\frac{1}{\sqrt{\theta}\sigma_1\sigma_2}d\theta d\sigma_1 d\sigma_2}\right\} \\
&\propto \exp\left\{\frac{\sqrt{b_j}\log b_j - \sqrt{a_j}\log a_j}{2(\sqrt{b_j}-\sqrt{a_j})} - \frac{\int_{1/\sqrt{2j}}^{\sqrt{2j}}\int_{1/\sqrt{j}}^{\sqrt{j}}\frac{\log(\omega^4\sigma_1^2+\sigma_2^2)}{\sigma_1\sigma_2}d\sigma_1 d\sigma_2}{2\log j\log(2j)}\right\}.
\end{aligned}$$

Set $\omega^0 = 1$. Then

$$\begin{aligned}
\pi_2^R(\omega) &= \lim_{j\to\infty}\frac{\pi_{2j}(\omega)}{\pi_{2j}(\omega^0)} = \lim_{j\to\infty}\frac{\pi_{2j}(\omega)}{\pi_{2j}(1)} \\
&\propto \lim_{j\to\infty}\exp\left\{\frac{\int_{1/\sqrt{2j}}^{\sqrt{2j}}\int_{1/\sqrt{j}}^{\sqrt{j}}\frac{\log(\sigma_1^2+\sigma_2^2)}{\sigma_1\sigma_2}d\sigma_1 d\sigma_2 - \int_{1/\sqrt{2j}}^{\sqrt{2j}}\int_{1/\sqrt{j}}^{\sqrt{j}}\frac{\log(\omega^4\sigma_1^2+\sigma_2^2)}{\sigma_1\sigma_2}d\sigma_1 d\sigma_2}{2\log j\log(2j)}\right\}.
\end{aligned}$$

By using several transformations and Taylor expansions in the integration,

$$\begin{aligned}
\int_{1/\sqrt{2j}}^{\sqrt{2j}}\int_{1/\sqrt{j}}^{\sqrt{j}}\frac{\log(\omega^4\sigma_1^2+\sigma_2^2)}{\sigma_1\sigma_2}d\sigma_1 d\sigma_2 &\approx \frac{\log(2j)\log^2(\omega^4 j)}{4} + \left(\frac{1}{8j^2}-\frac{1}{2}\right)\left(\omega^4+\frac{1}{\omega^4}\right) + \frac{\log^3(2j)}{12}, \\
\int_{1/\sqrt{2j}}^{\sqrt{2j}}\int_{1/\sqrt{j}}^{\sqrt{j}}\frac{\log(\sigma_1^2+\sigma_2^2)}{\sigma_1\sigma_2}d\sigma_1 d\sigma_2 &\approx \frac{\log(2j)\log^2 j}{4} + \frac{1}{4j^2} - 1 + \frac{\log^3(2j)}{12}.
\end{aligned}$$

Thus

$$\pi_2^R(\omega) \propto \lim_{j\to\infty} \exp\left\{\frac{B_{1j}+B_{2j}}{2\log j \log(2j)}\right\},$$

where

$$\begin{aligned} B_{1j} &= \frac{\log(2j)\{\log^2 j - \log^2(\omega^4 j)\}}{4} = -2\log(2j)(\log j \log\omega + 2\log^2\omega), \\ B_{2j} &= \left(\frac{1}{8j^2} - \frac{1}{2}\right)\left(2 - \omega^4 - \frac{1}{\omega^4}\right). \end{aligned}$$

It is easily proven that

$$\begin{aligned} \lim_{j\to\infty} \frac{B_{1j}}{2\log j \log(2j)} &= -\log\omega, \\ \lim_{j\to\infty} \frac{B_{2j}}{2\log j \log(2j)} &= 0. \end{aligned}$$

Consequently,

$$\pi_2^R(\omega) \propto \exp(-\log\omega) = \frac{1}{\omega}.$$

The result holds. $\square$

By Sun and Ye (1999), the reference prior for $(\theta, \omega, \sigma_1, \sigma_2)$ in the grouped ordering of $\{\theta, (\omega, \sigma_1, \sigma_2)\}$, where θ is the interest parameter and $(\omega, \sigma_1, \sigma_2)$ is the group of nuisance parameters with the same importance, is expressed as

$$\pi(\theta, \omega, \sigma_1, \sigma_2) \propto \frac{g(\theta)}{\sqrt{\theta}\sigma_1^2\sigma_2^2}\sqrt{\sigma_1^2 + \frac{\sigma_2^2}{\omega^4}},$$

where $g(\theta)$ is any positive real function. Previously, Berger and Bernardo (1989) computed the reference prior for (θ, ω) where σ_1 and σ_2 are known, when the parameter of interest is θ and nuisance parameter is ω. Sun and Ye (1995) extended it by considering more normal means.

3.9.2 Equal Variances

Assume $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Let

$$\theta_1 = \frac{\mu_1 - \mu_2}{\sigma} \quad \text{and} \quad \theta_2 = \frac{\mu_1 + \mu_2}{\sigma}.$$

The Fisher information matrix of $(\theta_1, \theta_2, \sigma)$ is

$$\mathbf{\Sigma}(\theta_1, \theta_2, \sigma) = \frac{1}{2} \begin{pmatrix} 1 & 0 & \frac{\theta_1}{\sigma} \\ 0 & 1 & \frac{\theta_2}{\sigma} \\ \frac{\theta_1}{\sigma} & \frac{\theta_2}{\sigma} & \frac{\theta_1^2+\theta_2^2+8}{\sigma^2} \end{pmatrix}. \tag{3.100}$$

It is easy to show

$$\begin{aligned} |\mathbf{\Sigma}| &= \frac{1}{\sigma^2}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{\theta_1^2+8}{4\sigma^2}, \quad |\mathbf{\Sigma}_{22}^c| = \frac{\theta_2^2+8}{4\sigma^2}, \quad |\mathbf{\Sigma}_{33}^c| = \frac{1}{4}, \\ |\mathbf{\Sigma}_{12}^c| &= |\mathbf{\Sigma}_{21}^c| = -\frac{\theta_1\theta_2}{4\sigma^2}, \quad |\mathbf{\Sigma}_{13}^c| = |\mathbf{\Sigma}_{31}^c| = -\frac{\theta_1}{4\sigma}, \quad |\mathbf{\Sigma}_{23}^c| = |\mathbf{\Sigma}_{32}^c| = \frac{\theta_2}{4\sigma}. \end{aligned}$$

Then

$$\begin{aligned} \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{4}{\theta_1^2+8}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} = \frac{4}{\theta_2^2+8}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} &= \frac{4}{\sigma^2}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} = \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = -\frac{4}{\theta_1\theta_2}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{13}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{31}^c|} = -\frac{4}{\theta_1\sigma}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{23}^c|} = \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{32}^c|} = \frac{4}{\theta_2\sigma}. \end{aligned}$$

Hence the marginal reference priors for θ_1, θ_2 and σ are

$$\pi_1^R(\theta_1) \propto \frac{1}{\sqrt{\theta_1^2+8}}, \; \theta_1 \in \mathbb{R}, \tag{3.101}$$

$$\pi_2^R(\theta_2) \propto \frac{1}{\sqrt{\theta_2^2+8}}, \; \theta_2 \in \mathbb{R}, \tag{3.102}$$

$$\pi_3^R(\sigma) \propto \frac{1}{\sigma}, \; \sigma > 0, \tag{3.103}$$

and the independent reference prior for $(\theta_1, \theta_2, \sigma)$ is

$$\pi^R(\theta_1, \theta_2, \sigma) \propto \frac{1}{\sigma\sqrt{(\theta_1^2+8)(\theta_2^2+8)}}. \tag{3.104}$$

3.9.3 Behrens-Fisher Problem

We are intrerested in $\theta = \mu_1 - \mu_2$, the difference between two means, under the unequal variances. We define $\omega = \mu_1 + \mu_2$, the sum of two means. It is easy to show that the Fisher information matrix of $(\theta, \omega, \sigma_1, \sigma_2)$ is

$$\mathbf{\Sigma}(\theta, \omega, \sigma_1, \sigma_2) = \begin{pmatrix} \frac{\sigma_1^2+\sigma_2^2}{4\sigma_1^2\sigma_2^2} & \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^2\sigma_2^2} & 0 & 0 \\ \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^2\sigma_2^2} & \frac{\sigma_1^2+\sigma_2^2}{4\sigma_1^2\sigma_2^2} & 0 & 0 \\ 0 & 0 & \frac{2}{\sigma_1^2} & 0 \\ 0 & 0 & 0 & \frac{2}{\sigma_2^2} \end{pmatrix}. \tag{3.105}$$

Thus

$$\begin{aligned} |\mathbf{\Sigma}| &= \frac{1}{\sigma_1^4\sigma_2^4}, \quad |\mathbf{\Sigma}_{11}^c| = |\mathbf{\Sigma}_{22}^c| = \frac{\sigma_1^2+\sigma_2^2}{\sigma_1^4\sigma_2^4}, \\ |\mathbf{\Sigma}_{33}^c| &= \frac{1}{2\sigma_1^2\sigma_2^4}, \quad |\mathbf{\Sigma}_{44}^c| = \frac{1}{2\sigma_1^4\sigma_2^2}, \\ |\mathbf{\Sigma}_{ij}^c| &= \begin{cases} \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^4\sigma_2^4}, & \text{if } (i,j) = (1,2), (2,1), \\ 0, & \text{otherwise}, \end{cases} \end{aligned}$$

and then

$$\begin{aligned} \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} = \frac{1}{\sigma_1^2+\sigma_2^2}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} &= \frac{2}{\sigma_1^2}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{44}^c|} = \frac{2}{\sigma_2^2}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = \frac{4}{\sigma_2^2-\sigma_1^2}. \end{aligned}$$

Hence the marginal reference priors for θ, ω, σ_1 and σ_2 are

$$\pi_1^R(\theta) \propto 1, \ \theta \in \mathbb{R}, \tag{3.106}$$

$$\pi_2^R(\omega) \propto 1, \ \omega \in \mathbb{R}, \tag{3.107}$$

$$\pi_3^R(\sigma_1) \propto \frac{1}{\sigma_1}, \ \sigma_1 > 0, \tag{3.108}$$

$$\pi_4^R(\sigma_2) \propto \frac{1}{\sigma_2}, \ \sigma_2 > 0, \tag{3.109}$$

and the independent reference prior for $(\theta, \omega, \sigma_1, \sigma_2)$ is

$$\pi^R(\theta, \omega, \sigma_1, \sigma_2) \propto \frac{1}{\sigma_1 \sigma_2}. \tag{3.110}$$

It is also a first order matching prior for θ, ω, σ_1 and σ_2, and the one-at-a-time reference prior for $(\theta, \omega, \sigma_1, \sigma_2)$ with any ordering.

3.9.4 Fieller-Creasy Problem

We are intrerested in the ratio of the two means, $\theta = \frac{\mu_1}{\mu_2}$, under the equal variances $\sigma_1^2 = \sigma_2^2 \equiv \sigma^2$. Then the Fisher information matrix of (θ, μ_2, σ) is

$$\mathbf{\Sigma}(\theta, \mu_2, \sigma) = \frac{1}{\sigma^2} \begin{pmatrix} \mu_2^2 & \theta\mu_2 & 0 \\ \theta\mu_2 & 1+\theta^2 & 0 \\ 0 & 0 & 4 \end{pmatrix}. \tag{3.111}$$

Thus

$$\begin{aligned} |\mathbf{\Sigma}| &= \frac{4\mu_2^2}{\sigma^6}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{4(1+\theta^2)}{\sigma^4}, \quad |\mathbf{\Sigma}_{22}^c| = \frac{4\mu_2^2}{\sigma^4}, \quad |\mathbf{\Sigma}_{33}^c| = \frac{\mu_2^2}{\sigma^4}, \\ |\mathbf{\Sigma}_{ij}^c| &= \begin{cases} \frac{4\theta\mu_2}{\sigma^4}, & \text{if } (i,j) = (1,2), (2,1), \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

and then

$$\begin{aligned} \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{\mu_2^2}{\sigma^2}\left(\frac{1}{1+\theta^2}\right), \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} &= \frac{1}{\sigma^2}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} = \frac{4}{\sigma^2}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = \frac{\mu_2}{\theta\sigma^2}. \end{aligned}$$

Hence the marginal reference priors for θ, μ_2 and σ are

$$\pi_1^R(\theta) \propto \frac{1}{\sqrt{1+\theta^2}}, \; \theta \in \mathbb{R}, \tag{3.112}$$

$$\pi_2^R(\mu_2) \quad \propto \quad 1, \ \mu_2 \in \mathbb{R}, \tag{3.113}$$

$$\pi_3^R(\sigma) \quad \propto \quad \frac{1}{\sigma}, \ \sigma > 0, \tag{3.114}$$

and the independent reference prior for (θ, μ_2, σ) is

$$\pi^R(\theta, \mu_2, \sigma) \propto \frac{1}{\sigma\sqrt{1+\theta^2}}. \tag{3.115}$$

It is a first order matching prior for θ and σ. The independent reference prior for (θ, μ_2, σ) is identical to the one-at-a-time reference prior for (θ, μ_2, σ) which was derived by Bernardo (1977), when θ is of interest.

3.10 Bivariate Normal Model

Let $(X_1, X_2)'$ be a bivariate normal random vector with unknown mean parameters $(\mu_1, \mu_2)'$ and unknown covariance matrix $\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$, whose density is given by

$$\begin{aligned} & f(x_1, x_2 \mid \mu_1, \mu_2, \sigma_1, \sigma_2, \rho) \\ = \ & \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{\sigma_2^2(x_1-\mu_1)^2 + \sigma_1^2(x_2-\mu_2)^2 - 2\rho\sigma_1\sigma_2(x_1-\mu_1)(x_2-\mu_2)}{2\sigma_1^2\sigma_2^2(1-\rho^2)}\right\}. \end{aligned}$$

Here $\rho \in (-1, 1)$ is the correlation between X_1 and X_2, and $x_i \in \mathbb{R}, \mu_i \in \mathbb{R}, \sigma_i > 0$ for $i = 1, 2$.

All the reparameterizations in this section were considered by Berger and Sun (2007). The Fisher information matrices and the reference priors referred here were also derived by Berger and Sun (2007).

3.10.1 Commonly Used Parameters

The Fisher information matrix of $(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ is

$$\mathbf{\Sigma}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho) = \frac{1}{1-\rho^2}\begin{pmatrix} \frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2} & 0 & 0 & 0 \\ -\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} & 0 & 0 & 0 \\ 0 & 0 & \frac{2-\rho^2}{\sigma_1^2} & -\frac{\rho^2}{\sigma_1\sigma_2} & -\frac{\rho}{\sigma_1} \\ 0 & 0 & -\frac{\rho^2}{\sigma_1\sigma_2} & \frac{2-\rho^2}{\sigma_2^2} & -\frac{\rho}{\sigma_2} \\ 0 & 0 & -\frac{\rho}{\sigma_1} & -\frac{\rho}{\sigma_2} & \frac{1+\rho^2}{1-\rho^2} \end{pmatrix}. \tag{3.116}$$

Thus

$$\begin{aligned}
|\mathbf{\Sigma}| &= \frac{4}{\sigma_1^4\sigma_2^4(1-\rho^2)^4}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{4}{\sigma_1^2\sigma_2^4(1-\rho^2)^4}, \quad |\mathbf{\Sigma}_{22}^c| = \frac{4}{\sigma_1^4\sigma_2^2(1-\rho^2)^4}, \\
|\mathbf{\Sigma}_{33}^c| &= \frac{2}{\sigma_1^2\sigma_2^4(1-\rho^2)^4}, \quad |\mathbf{\Sigma}_{44}^c| = \frac{2}{\sigma_1^4\sigma_2^2(1-\rho^2)^4}, \quad |\mathbf{\Sigma}_{55}^c| = \frac{4}{\sigma_1^4\sigma_2^4(1-\rho^2)^2}, \\
|\mathbf{\Sigma}_{12}^c| &= |\mathbf{\Sigma}_{21}^c| = -\frac{4\rho}{\sigma_1^3\sigma_2^3(1-\rho^2)^4}, \quad |\mathbf{\Sigma}_{34}^c| = |\mathbf{\Sigma}_{43}^c| = -\frac{2\rho^2}{\sigma_1^3\sigma_2^3(1-\rho^2)^4}, \\
|\mathbf{\Sigma}_{35}^c| &= |\mathbf{\Sigma}_{53}^c| = \frac{2\rho}{\sigma_1^3\sigma_2^4(1-\rho^2)^3}, \quad |\mathbf{\Sigma}_{45}^c| = |\mathbf{\Sigma}_{54}^c| = -\frac{2\rho}{\sigma_1^4\sigma_2^3(1-\rho^2)^3}, \\
|\mathbf{\Sigma}_{13}^c| &= |\mathbf{\Sigma}_{31}^c| = |\mathbf{\Sigma}_{14}^c| = |\mathbf{\Sigma}_{41}^c| = |\mathbf{\Sigma}_{15}^c| = |\mathbf{\Sigma}_{51}^c| = 0, \\
|\mathbf{\Sigma}_{23}^c| &= |\mathbf{\Sigma}_{32}^c| = |\mathbf{\Sigma}_{24}^c| = |\mathbf{\Sigma}_{42}^c| = |\mathbf{\Sigma}_{25}^c| = |\mathbf{\Sigma}_{52}^c| = 0,
\end{aligned}$$

and then

$$\begin{aligned}
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{1}{\sigma_1^2}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} = \frac{1}{\sigma_2^2}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} &= \frac{2}{\sigma_1^2}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{44}^c|} = \frac{2}{\sigma_2^2}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{55}^c|} &= \frac{1}{(1-\rho^2)^2}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = -\frac{1}{\sigma_1\sigma_2\rho}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{34}^c|} = \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{43}^c|} = -\frac{2}{\sigma_1\sigma_2\rho^2}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{35}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{53}^c|} = \frac{2}{\sigma_1\rho(1-\rho^2)}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{45}^c|} = \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{54}^c|} = -\frac{2}{\sigma_2\rho(1-\rho^2)}.
\end{aligned}$$

Hence the marginal reference priors for μ_1, μ_2, σ_1, σ_2 and ρ are

$$\begin{aligned} \pi_1^R(\mu_1) &\propto 1,\ \mu_1 \in \mathbb{R}, && (3.117)\\ \pi_2^R(\mu_2) &\propto 1,\ \mu_2 \in \mathbb{R}, && (3.118)\\ \pi_3^R(\sigma_1) &\propto \frac{1}{\sigma_1},\ \sigma_1 > 0, && (3.119)\\ \pi_4^R(\sigma_2) &\propto \frac{1}{\sigma_2},\ \sigma_2 > 0, && (3.120)\\ \pi_5^R(\rho) &\propto \frac{1}{1-\rho^2},\ \rho \in (-1,1), && (3.121) \end{aligned}$$

and the independent reference prior for $(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ is

$$\pi^R(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho) \propto \frac{1}{\sigma_1\sigma_2(1-\rho^2)}. \qquad (3.122)$$

It is a first order matching prior for μ_1, μ_2 and ρ, and the one-at-a-time reference prior for $(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ in the ordering of $\{\rho, \sigma_1, \sigma_2, \mu_1, \mu_2\}$, $\{\rho, \sigma_2, \sigma_1, \mu_1, \mu_2\}$, $\{\mu_1, \mu_2, \rho, \sigma_1, \sigma_2\}$ and $\{\mu_1, \mu_2, \rho, \sigma_2, \sigma_1\}$. Note that Berger and Sun (2007) also derived the Jeffreys-rule prior, π^J and the independence Jeffreys prior, π^{IJ}. They are given by

$$\begin{aligned} \pi^J(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho) &= \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}, && (3.123)\\ \pi^{IJ}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho) &= \frac{1}{\sigma_1\sigma_2(1-\rho^2)^{3/2}}. && (3.124) \end{aligned}$$

Any of the Jeffreys priors are not the same as the independent reference prior given by (3.122).

We can consider alternative reparameterization, $\theta = \frac{\sigma_2^2}{\sigma_1^2}$ and $\xi = \sigma_1\sigma_2$. Then the Fisher information matrix of $(\mu_1, \mu_2, \theta, \rho, \xi)$ is

$$\mathbf{\Sigma}(\mu_1, \mu_2, \theta, \rho, \xi) = \frac{1}{1-\rho^2} \begin{pmatrix} \frac{\sqrt{\theta}}{\xi} & -\frac{\rho}{\xi} & 0 & 0 & 0 \\ -\frac{\rho}{\xi} & \frac{1}{\xi\sqrt{\theta}} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\theta^2} & 0 & 0 \\ 0 & 0 & 0 & 1+\rho^2 & -\frac{\rho}{\xi} \\ 0 & 0 & 0 & -\frac{\rho}{\xi} & \frac{1}{\xi^2} \end{pmatrix}. \tag{3.125}$$

Thus

$$\begin{aligned}
|\mathbf{\Sigma}| &= \frac{1}{\theta^2\xi^4(1-\rho^2)^4}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{1}{\theta^{5/2}\xi^3(1-\rho^2)^4}, \\
|\mathbf{\Sigma}_{22}^c| &= \frac{1}{\theta^{3/2}\xi^3(1-\rho^2)^4}, \quad |\mathbf{\Sigma}_{33}^c| = \frac{1}{\xi^4(1-\rho^2)^3}, \\
|\mathbf{\Sigma}_{44}^c| &= \frac{1}{\theta^2\xi^4(1-\rho^2)^2}, \quad |\mathbf{\Sigma}_{55}^c| = \frac{1+\rho^2}{\theta^2\xi^2(1-\rho^2)^3}, \\
|\mathbf{\Sigma}_{ij}^c| &= \begin{cases} -\frac{\rho}{\theta^2\xi^3(1-\rho^2)^4}, & \text{if } (i,j) = (1,2), (2,1), \\ -\frac{\rho}{\theta^2\xi^3(1-\rho^2)^3}, & \text{if } (i,j) = (4,5), (5,4), \\ 0, & \text{otherwise,} \end{cases}
\end{aligned}$$

and then

$$\begin{aligned}
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{\sqrt{\theta}}{\xi}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} = \frac{1}{\xi\sqrt{\theta}}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} &= \frac{1}{\theta^2(1-\rho^2)}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{44}^c|} = \frac{1}{(1-\rho^2)^2}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{55}^c|} = \frac{1}{\xi^2(1-\rho^4)}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = -\frac{1}{\xi\rho}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{45}^c|} = \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{54}^c|} = -\frac{1}{\xi\rho(1-\rho^2)}.
\end{aligned}$$

Hence the marginal reference priors for μ_1, μ_2, θ, ρ and ξ are

$$\pi_1^R(\mu_1) \propto 1, \ \mu_1 \in \mathbb{R}, \tag{3.126}$$

$$\pi_2^R(\mu_2) \propto 1, \ \mu_2 \in \mathbb{R}, \tag{3.127}$$

$$\pi_3^R(\theta) \propto \frac{1}{\theta}, \ \theta > 0, \tag{3.128}$$

$$\pi_4^R(\rho) \propto \frac{1}{1-\rho^2},\ \rho \in (-1,1), \tag{3.129}$$

$$\pi_5^R(\xi) \propto \frac{1}{\xi},\ \xi > 0, \tag{3.130}$$

and the independent reference prior for $(\mu_1, \mu_2, \theta, \rho, \xi)$ is

$$\pi^R(\mu_1, \mu_2, \theta, \rho, \xi) \propto \frac{1}{\theta\xi(1-\rho^2)}. \tag{3.131}$$

It is is a first order matching prior for μ_1, μ_2, θ and ρ, and also the one-at-a-time reference prior for $(\mu_1, \mu_2, \theta, \rho, \xi)$ in the ordering of $\{\theta, \rho, \xi, \mu_1, \mu_2\}$ and $\{\rho, \theta, \xi, \mu_1, \mu_2\}$.

3.10.2 Cholesky Decomposition

Define

$$\eta_1 = \frac{1}{\sigma_1},\quad \eta_2 = \frac{1}{\sigma_2\sqrt{1-\rho^2}} \text{ and } \eta_3 = -\frac{\rho}{\sigma_1\sqrt{1-\rho^2}}. \tag{3.132}$$

It is easy to verify that

$$\mathbf{\Omega}^{-1} = \begin{pmatrix} \eta_1 & \eta_3 \\ 0 & \eta_2 \end{pmatrix} \begin{pmatrix} \eta_1 & 0 \\ \eta_3 & \eta_2 \end{pmatrix}.$$

So (η_1, η_2, η_3) is a set of parameters for a type of Cholesky decomposition of $\mathbf{\Omega}^{-1}$. The Fisher information matrix of $(\mu_1, \mu_2, \eta_1, \eta_2, \eta_3)$ is

$$\mathbf{\Sigma}(\mu_1, \mu_2, \eta_1, \eta_2, \eta_3) = \begin{pmatrix} \eta_1^2+\eta_3^2 & \eta_2\eta_3 & 0 & 0 & 0 \\ \eta_2\eta_3 & \eta_2^2 & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{\eta_1^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{2\eta_1^2+\eta_3^2}{\eta_1^2\eta_2^2} & -\frac{\eta_3}{\eta_1^2\eta_2} \\ 0 & 0 & 0 & -\frac{\eta_3}{\eta_1^2\eta_2} & \frac{1}{\eta_1^2} \end{pmatrix}. \tag{3.133}$$

Thus

$$
\begin{aligned}
|\boldsymbol{\Sigma}| &= \frac{4}{\eta_1^2}, \quad |\boldsymbol{\Sigma}_{11}^c| = \frac{4}{\eta_1^4}, \quad |\boldsymbol{\Sigma}_{22}^c| = \frac{4(\eta_1^2+\eta_3^2)}{\eta_1^4\eta_2^2}, \\
|\boldsymbol{\Sigma}_{33}^c| &= 2, \quad |\boldsymbol{\Sigma}_{44}^c| = \frac{2\eta_2^2}{\eta_1^2}, \quad |\boldsymbol{\Sigma}_{55}^c| = \frac{4(\eta_1^2+\eta_3^2/2)}{\eta_1^2},
\end{aligned}
$$

and then

$$
\begin{aligned}
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}^c|} &= \eta_1^2, \quad \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{22}^c|} = \frac{\eta_1^2\eta_2^2}{\eta_1^2+\eta_3^2}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{33}^c|} &= \frac{2}{\eta_1^2}, \quad \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{44}^c|} = \frac{2}{\eta_2^2}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{55}^c|} &= \frac{1}{\eta_1^2+\eta_3^2/2}.
\end{aligned}
$$

It is easily seen that $|\boldsymbol{\Sigma}|/|\boldsymbol{\Sigma}_{ii}^c|$, $i = 1, \ldots, 4$, satisfy Condition (2.3) but $|\boldsymbol{\Sigma}|/|\boldsymbol{\Sigma}_{55}^c|$ does not. Thus we cannot apply Theorem 2.1 to this case. Now, we return to the iterative algorithm and use it to compute the independent reference prior for $(\mu_1, \mu_2, \eta_1, \eta_2, \eta_3)$.

Proposition 3.6 *Suppose that* $\{[1/j, j^j],\ j = 1, 2, \ldots\}$ *is an increasing sequence of compact subsets of* $(0, \infty)$ *for* η_1. *Then the marginal reference priors for* μ_1, μ_2, η_1, η_2 *and* η_3 *are*

$$\pi_1^R(\mu_1) \propto 1, \ \mu_1 \in \mathbb{R}, \tag{3.134}$$

$$\pi_2^R(\mu_2) \propto 1, \ \mu_2 \in \mathbb{R}, \tag{3.135}$$

$$\pi_3^R(\eta_1) \propto \frac{1}{\eta_1}, \ \eta_1 > 0, \tag{3.136}$$

$$\pi_4^R(\eta_2) \propto \frac{1}{\eta_2}, \ \eta_2 > 0, \tag{3.137}$$

$$\pi_5^R(\eta_3) \propto 1, \ \eta_3 \in \mathbb{R}. \tag{3.138}$$

Consequently, the independent reference prior for $(\mu_1, \mu_2, \eta_1, \eta_2, \eta_3)$ *is*

$$\pi^R(\mu_1, \mu_2, \eta_1, \eta_2, \eta_3) \propto \frac{1}{\eta_1\eta_2}. \tag{3.139}$$

Proof. It is shown that $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{ii}^c|$, $i = 1, \ldots, 4$, satisfy Condition (2.3) so that (3.134)–(3.137) hold immediately. It is easily seen that $\pi_1^R(\mu_1)$, $\pi_2^R(\mu_2)$, $\pi_3^R(\eta_1)$ and $\pi_4^R(\eta_2)$ are improper. Thus we need an argument of compact sets and use *Algorithm B′* to derive $\pi_5^R(\eta_3)$ since $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{55}^c|$ does not meet Condition (2.3). Choose $\{[a_{ij}, b_{ij}],\ j = 1, 2, \ldots\}$, $i = 1, 2$, as an increasing sequence of compact subsets of $\mathbb{R}$ for μ_1 and μ_2 respectively, where $a_{ij} \to -\infty$ and $b_{ij} \to \infty$ as $j \to \infty$ for $i = 1, 2$. Also choose $\{[1/j, j^j],\ j = 1, 2, \ldots\}$ as an increasing sequence of compact subsets of $(0, \infty)$ for η_1 and $\{[a_{3j}, b_{3j}],\ j = 1, 2, \ldots\}$ for η_2, where $a_{3j} \to 0$ and $b_{3j} \to \infty$ as $j \to \infty$. Then by applying *Algorithm B′*, it can be shown that

$$\begin{aligned}
\pi_{5j}(\eta_3) &\propto \exp\left\{\frac{\int_{a_{3j}}^{b_{3j}} \int_{1/j}^{j^j} \int_{a_{2j}}^{b_{2j}} \int_{a_{1j}}^{b_{1j}} \pi_1^R(\mu_1)\pi_2^R(\mu_2)\pi_3^R(\eta_1)\pi_4^R(\eta_2) \log\left(\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{55}^c|}\right) d\mu_1 d\mu_2 d\eta_1 d\eta_2}{-2\int_{a_{3j}}^{b_{3j}} \int_{1/j}^{j^j} \int_{a_{2j}}^{b_{2j}} \int_{a_{1j}}^{b_{1j}} \pi_1^R(\mu_1)\pi_2^R(\mu_2)\pi_3^R(\eta_1)\pi_4^R(\eta_2) d\mu_1 d\mu_2 d\eta_1 d\eta_2}\right\} \\
&= \exp\left\{\frac{\int_{a_{3j}}^{b_{3j}} \int_{1/j}^{j^j} \frac{1}{\eta_1\eta_2} \log\left(\eta_1^2 + \frac{\eta_3^2}{2}\right) d\eta_1 d\eta_2}{-2\int_{a_{3j}}^{b_{3j}} \int_{1/j}^{j^j} \frac{1}{\eta_1\eta_2} d\eta_1 d\eta_2}\right\} \\
&= \exp\left\{\frac{\int_{1/j}^{j^j} \frac{1}{\eta_1} \log\left(\eta_1^2 + \frac{\eta_3^2}{2}\right) d\eta_1}{-2(j+1)\log j}\right\}.
\end{aligned}$$

Set $\eta_3^0 = 0$. Then

$$\begin{aligned}
\pi_5^R(\eta_3) &= \lim_{j\to\infty} \frac{\pi_{5j}(\eta_3)}{\pi_{5j}(\eta_3^0)} = \lim_{j\to\infty} \frac{\pi_{5j}(\eta_3)}{\pi_{5j}(0)} \\
&\propto \lim_{j\to\infty} \exp\left\{\frac{2\int_{1/j}^{j^j} \frac{1}{\eta_1} \log \eta_1 d\eta_1 - \int_{1/j}^{j^j} \frac{1}{\eta_1} \log\left(\eta_1^2 + \frac{\eta_3^2}{2}\right) d\eta_1}{2(j+1)\log j}\right\} \\
&= \lim_{j\to\infty} \exp\left\{\frac{(j-1)\log j}{2} - \frac{\int_{1/j}^{j^j} \frac{1}{\eta_1} \log\left(\eta_1^2 + \frac{\eta_3^2}{2}\right) d\eta_1}{2(j+1)\log j}\right\}.
\end{aligned}$$

By using several transformations and Taylor expansions in the integration, we obtain

$$\int_{1/j}^{j^j} \frac{1}{\eta_1} \log\left(\eta_1^2 + \frac{\eta_3^2}{2}\right) d\eta_1 \approx j^2 \log^2 j + \log j \log\left(\frac{\eta_3^2}{2}\right) - \frac{2}{j^2\eta_3^2} + \frac{\eta_3^2}{4j^{2j}} + \frac{\log^2(\eta_3^2/2)}{4}.$$

Thus

$$\pi_5^R(\eta_3) \propto \lim_{j\to\infty} \exp\left\{-\frac{B_{1j} + B_{2j}}{2(j+1)\log j}\right\},$$

where

$$B_{1j} = \log^2 j \quad \text{and} \quad B_{2j} = \log j \log\left(\frac{\eta_3^2}{2}\right) - \frac{2}{j^2\eta_3^2} + \frac{\eta_3^2}{4j^{2j}} + \frac{\log^2(\eta_3^2/2)}{4}.$$

It is easily shown that since $\lim_{j\to\infty} \log j/(j+1) = 0$,

$$\lim_{j\to\infty} \frac{B_{1j}}{2(j+1)\log j} = \lim_{j\to\infty} \frac{B_{2j}}{2(j+1)\log j} = 0.$$

Hence

$$\pi_5^R(\eta_3) \quad \propto \quad \exp(0) = 1.$$

The result then holds. □

Berger and Sun (2007) derived the reference priors for $(\mu_1, \mu_2, \eta_1, \eta_2, \eta_3)$ in the ordering of $\{\mu_1, \mu_2, \eta_1, \eta_2, \eta_3\}$, $\{\mu_1, \mu_2, \eta_1, \eta_3, \eta_2\}$ and $\{\mu_1, \mu_2, \eta_1, (\eta_2, \eta_3)\}$. The reference prior for the ordering of $\{\mu_1, \mu_2, \eta_1, \eta_2, \eta_3\}$ and $\{\mu_1, \mu_2, \eta_1, (\eta_2, \eta_3)\}$ is

$$\pi(\mu_1, \mu_2, \eta_1, \eta_2, \eta_3) \propto \frac{1}{\eta_1\eta_2},$$

which is the same as the independent reference prior for $(\mu_1, \mu_2, \eta_1, \eta_2, \eta_3)$ given by (3.139). For the ordering of $\{\mu_1, \mu_2, \eta_1, \eta_3, \eta_2\}$, the reference prior is

$$\pi(\mu_1, \mu_2, \eta_1, \eta_2, \eta_3) \propto \frac{1}{\eta_1\eta_2\sqrt{\eta_1^2 + \eta_3^2/2}}.$$

3.10.3 Orthogonal Parameterizations

Define

$$\omega = \frac{\eta_3}{\eta_2} = -\frac{\rho\sigma_2}{\sigma_1}. \tag{3.140}$$

It is easy to show that

$$\mathbf{\Omega}^{-1} = \begin{pmatrix} 1 & \omega \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_1^2 & 0 \\ 0 & \eta_2^2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \omega & 1 \end{pmatrix} = \begin{pmatrix} \eta_1 & \eta_2\omega \\ 0 & \eta_2 \end{pmatrix} \begin{pmatrix} \eta_1 & 0 \\ \eta_2\omega & \eta_2 \end{pmatrix}$$

since

$$\begin{pmatrix} \eta_1 & \eta_3 \\ 0 & \eta_2 \end{pmatrix} = \begin{pmatrix} 1 & \omega \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{pmatrix}.$$

Thus (η_1, η_2, ω) is a set of parameters for a type of Cholesky decomposition of $\mathbf{\Omega}^{-1}$. The Fisher information matrix of $(\mu_1, \mu_2, \eta_1, \eta_2, \omega)$ is

$$\mathbf{\Sigma}(\mu_1, \mu_2, \eta_1, \eta_2, \omega) = \begin{pmatrix} \eta_1^2 + \eta_2^2\omega^2 & \eta_2^2\omega & 0 & 0 & 0 \\ \eta_2^2\omega & \eta_2^2 & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{\eta_1^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{2}{\eta_2^2} & 0 \\ 0 & 0 & 0 & 0 & \frac{\eta_2^2}{\eta_1^2} \end{pmatrix}. \tag{3.141}$$

Note that the Fisher information matrix is almost diagonal, except for the left-top corner corresponding to (μ_1, μ_2). Thus

$$\begin{aligned} |\mathbf{\Sigma}| &= \frac{4\eta_2^2}{\eta_1^2}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{4\eta_2^2}{\eta_1^4}, \quad |\mathbf{\Sigma}_{22}^c| = \frac{4(\eta_1^2 + \eta_2^2\omega^2)}{\eta_1^4}, \\ |\mathbf{\Sigma}_{33}^c| &= 2\eta_2^2, \quad |\mathbf{\Sigma}_{44}^c| = \frac{2\eta_2^4}{\eta_1^2}, \quad |\mathbf{\Sigma}_{55}^c| = 4, \\ |\mathbf{\Sigma}_{ij}^c| &= \begin{cases} \frac{4\eta_2^2\omega}{\eta_1^4}, & \text{if } (i,j) = (1,2), (2,1), \\ 0, & \text{otherwise}, \end{cases} \end{aligned}$$

and then

$$\begin{aligned} \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \eta_1^2, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} = \frac{\eta_1^2\eta_2^2}{\eta_1^2 + \eta_2^2\omega^2}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} &= \frac{2}{\eta_1^2}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{44}^c|} = \frac{2}{\eta_2^2}, \quad \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{55}^c|} = \frac{\eta_2^2}{\eta_1^2}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = \frac{\eta_1^2}{\omega}. \end{aligned}$$

Hence the marginal reference priors for μ_1, μ_2, η_1, η_2 and ω are

$$\pi_1^R(\mu_1) \propto 1, \ \mu_1 \in \mathbb{R}, \tag{3.142}$$

$$\pi_2^R(\mu_2) \propto 1, \ \mu_2 \in \mathbb{R}, \tag{3.143}$$

$$\pi_3^R(\eta_1) \propto \frac{1}{\eta_1}, \ \eta_1 > 0, \tag{3.144}$$

$$\pi_4^R(\eta_2) \propto \frac{1}{\eta_2}, \ \eta_2 > 0, \tag{3.145}$$

$$\pi_5^R(\omega) \propto 1, \ \omega \in \mathbb{R}, \tag{3.146}$$

and the independent reference prior for $(\mu_1, \mu_2, \eta_1, \eta_2, \omega)$ is

$$\pi^R(\mu_1, \mu_2, \eta_1, \eta_2, \omega) \propto \frac{1}{\eta_1\eta_2}. \tag{3.147}$$

It is also a first order matching prior for μ_1, μ_2, η_1, η_2 and ω, and the one-at-a-time reference prior for $(\mu_1, \mu_2, \eta_1, \eta_2, \omega)$ with any ordering.

Alternatively, define

$$\xi_1 = \eta_1\eta_2 = \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}} \quad \text{and} \quad \xi_2 = \frac{\eta_1}{\eta_2} = \frac{\sigma_2\sqrt{1-\rho^2}}{\sigma_1}. \tag{3.148}$$

Then the Fisher information matrix of $(\mu_1, \mu_2, \xi_1, \xi_2, \omega)$ is

$$\mathbf{\Sigma}(\mu_1, \mu_2, \xi_1, \xi_2, \omega) = \begin{pmatrix} \frac{\xi_1(\xi_2^2+\omega^2)}{\xi_2} & \frac{\omega\xi_1}{\xi_2} & 0 & 0 & 0 \\ \frac{\omega\xi_1}{\xi_2} & \frac{\xi_1}{\xi_2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\xi_1^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\xi_2^2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\xi_2^2} \end{pmatrix}. \tag{3.149}$$

Note again that the Fisher information matrix is almost diagonal, except for the left-top corner corresponding to (μ_1, μ_2). Thus

$$|\mathbf{\Sigma}| = \frac{1}{\xi_2^4}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{1}{\xi_1\xi_2^5}, \quad |\mathbf{\Sigma}_{22}^c| = \frac{\xi_2^2+\omega^2}{\xi_1\xi_2^5},$$

$$
\begin{aligned}
|\boldsymbol{\Sigma}_{33}^c| &= \frac{\xi_1^2}{\xi_2^4}, \quad |\boldsymbol{\Sigma}_{44}^c| = |\boldsymbol{\Sigma}_{55}^c| = \frac{1}{\xi_2^2}, \\
|\boldsymbol{\Sigma}_{ij}^c| &= \begin{cases} \frac{\omega}{\xi_1\xi_2^5}, & \text{if } (i,j) = (1,2),(2,1), \\ 0, & \text{otherwise,} \end{cases}
\end{aligned}
$$

and then

$$
\begin{aligned}
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}^c|} &= \xi_1\xi_2, \quad \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{22}^c|} = \frac{\xi_1\xi_2}{\xi_2^2+\omega^2}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{33}^c|} &= \frac{1}{\xi_1^2}, \quad \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{44}^c|} = \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{55}^c|} = \frac{1}{\xi_2^2}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{12}^c|} &= \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{21}^c|} = \frac{\xi_1\xi_2}{\omega}.
\end{aligned}
$$

Hence the marginal reference priors for μ_1, μ_2, ξ_1, ξ_2 and ω are

$$\pi_1^R(\mu_1) \propto 1, \; \mu_1 \in \mathbb{R}, \tag{3.150}$$

$$\pi_2^R(\mu_2) \propto 1, \; \mu_2 \in \mathbb{R}, \tag{3.151}$$

$$\pi_3^R(\xi_1) \propto \frac{1}{\xi_1}, \; \xi_1 > 0, \tag{3.152}$$

$$\pi_4^R(\xi_2) \propto \frac{1}{\xi_2}, \; \xi_2 > 0, \tag{3.153}$$

$$\pi_5^R(\omega) \propto 1, \; \omega \in \mathbb{R}, \tag{3.154}$$

and the independent reference prior for $(\mu_1, \mu_2, \xi_1, \xi_2, \omega)$ is

$$\pi^R(\mu_1, \mu_2, \xi_1, \xi_2, \omega) \propto \frac{1}{\xi_1\xi_2}. \tag{3.155}$$

It is a first order matching prior for μ_1, μ_2, ξ_1, ξ_2 and ω, and also the one-at-a-time reference prior for $(\mu_1, \mu_2, \xi_1, \xi_2, \omega)$ with any ordering.

3.11 Poisson Model: Two Independent Samples

Let X_1 and X_2 be independent Poisson random variables with means λ_1 and λ_2. Then for $x_i \in \{0, 1, 2, \ldots\}, \lambda_i > 0$, $i = 1, 2$, the joint density of (X_1, X_2) is

$$f(x_1, x_2|\lambda_1, \lambda_2) = \frac{e^{-\lambda_1}\lambda_1^{x_1}}{x_1!}\frac{e^{-\lambda_2}\lambda_2^{x_2}}{x_2!}.$$

Then the Fisher information matrix of (λ_1, λ_2) is

$$\boldsymbol{\Sigma}(\lambda_1, \lambda_2) = \begin{pmatrix} \frac{1}{\lambda_1} & 0 \\ 0 & \frac{1}{\lambda_2} \end{pmatrix}. \tag{3.156}$$

Hence the marginal reference priors for λ_1 and λ_2 are

$$\pi_1^R(\lambda_1) \propto \frac{1}{\sqrt{\lambda_1}},\ \lambda_1 > 0, \tag{3.157}$$

$$\pi_2^R(\lambda_2) \propto \frac{1}{\sqrt{\lambda_2}},\ \lambda_2 > 0, \tag{3.158}$$

and the independent reference prior for (λ_1, λ_2) is

$$\pi^R(\lambda_1, \lambda_2) \propto \frac{1}{\sqrt{\lambda_1\lambda_2}}. \tag{3.159}$$

It is also a first order matching prior for λ_1 and λ_2, and the reference prior for (λ_1, λ_2) when one of the parameters λ_1 or λ_2 is the interest parameter and the other is nuisance parameter.

We consider alternative reparameterization, $\theta = \frac{\lambda_1}{\lambda_2}$, the ratio of two means. Then the Fisher information matrix of (θ, λ_2) is

$$\boldsymbol{\Sigma}(\theta, \lambda_2) = \begin{pmatrix} \frac{\lambda_2}{\theta} & 1 \\ 1 & \frac{1+\theta}{\lambda_2} \end{pmatrix}. \tag{3.160}$$

Thus

$$|\boldsymbol{\Sigma}| = \frac{1}{\theta},\quad |\boldsymbol{\Sigma}_{11}^c| = \frac{1+\theta}{\lambda_2},\quad |\boldsymbol{\Sigma}_{22}^c| = \frac{\lambda_2}{\theta},\quad |\boldsymbol{\Sigma}_{12}^c| = |\boldsymbol{\Sigma}_{21}^c| = 1,$$

and then

$$\begin{aligned}\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}^c|} &= \frac{\lambda_2}{\theta(1+\theta)},\\ \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{22}^c|} &= \frac{1}{\lambda_2},\\ \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{12}^c|} &= \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{21}^c|} = \frac{1}{\theta}.\end{aligned}$$

Hence the marginal reference priors for θ and λ_2 are

$$\pi_1^R(\theta) \propto \frac{1}{\sqrt{\theta(1+\theta)}}, \ \theta > 0, \tag{3.161}$$

$$\pi_2^R(\lambda_2) \propto \frac{1}{\sqrt{\lambda_2}}, \ \lambda_2 > 0, \tag{3.162}$$

and the independent reference prior for (θ, λ_2) is

$$\pi^R(\theta, \lambda_2) \propto \frac{1}{\sqrt{\lambda_2\theta(1+\theta)}}. \tag{3.163}$$

It is a first order matching prior for θ, and also the reference prior for (θ, λ_2) when θ is of interest and λ_2 is nuisance parameter.

3.12 Weibull Model

The Weibull density, denoted by $W(\eta, \beta)$, is

$$f(x|\eta, \beta) = \frac{\beta x^{\beta-1}}{\eta^\beta} \exp\left\{-\left(\frac{x}{\eta}\right)^\beta\right\}, \quad x > 0. \tag{3.164}$$

Here $\eta > 0$ is a unknown scale parameter and $\beta > 0$ is a unknown shape parameter. Then the Fisher information matrix of (η, β) is

$$\boldsymbol{\Sigma}(\eta, \beta) = \begin{pmatrix} \frac{\beta^2}{\eta^2} & -\frac{1+\gamma_1}{\eta} \\ -\frac{1+\gamma_1}{\eta} & \frac{\gamma_2+2\gamma_1+1}{\beta^2} \end{pmatrix}, \tag{3.165}$$

where $\gamma_i = \int_0^\infty [\log(z)]^i e^{-z} dz$. Thus

$$
\begin{aligned}
|\boldsymbol{\Sigma}| &= \frac{\gamma_2 - \gamma_1^2}{\eta^2}, \quad |\boldsymbol{\Sigma}_{11}^c| = \frac{\gamma_2 + 2\gamma_1 + 1}{\beta^2}, \\
|\boldsymbol{\Sigma}_{22}^c| &= \frac{\beta^2}{\eta^2}, \quad |\boldsymbol{\Sigma}_{12}^c| = |\boldsymbol{\Sigma}_{21}^c| = -\frac{1+\gamma_1}{\eta},
\end{aligned}
$$

and then

$$
\begin{aligned}
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}^c|} &= \frac{1}{\eta^2}\left\{\frac{\beta^2(\gamma_2 - \gamma_1^2)}{\gamma_2 + 2\gamma_1 + 1}\right\}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{22}^c|} &= \frac{\gamma_2 - \gamma_1^2}{\beta^2}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{12}^c|} &= \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{21}^c|} = -\frac{\gamma_2 - \gamma_1^2}{\eta(1+\gamma_1)}.
\end{aligned}
$$

Hence the marginal reference priors for η and β are

$$
\pi_1^R(\eta) \propto \frac{1}{\eta}, \ \eta > 0, \tag{3.166}
$$

$$
\pi_2^R(\beta) \propto \frac{1}{\beta}, \ \beta > 0, \tag{3.167}
$$

and the independent reference prior for (η, β) is

$$
\pi^R(\eta, \beta) \propto \frac{1}{\eta\beta}. \tag{3.168}
$$

It is also a first order matching prior for η and β, and the reference prior for (η, β) when one of the parameters η or β is the interest and the other is nuisance parameter.

There are three other Weibull densities which are given by

$$
f(x|\alpha, \beta) = \alpha^\beta \beta x^{\beta-1} \exp\left\{-(\alpha x)^\beta\right\}, \quad x > 0, \alpha > 0, \beta > 0, \tag{3.169}
$$

$$
f(x|\theta, \beta) = \frac{\beta x^{\beta-1}}{\theta} \exp\left(-\frac{x^\beta}{\theta}\right), \quad x > 0, \theta > 0, \beta > 0, \tag{3.170}
$$

$$
f(x|\lambda, \beta) = \lambda\beta x^{\beta-1} \exp\left(-\lambda x^\beta\right), \quad x > 0, \lambda > 0, \beta > 0. \tag{3.171}
$$

The parameters (η, β) in (3.164) and (α, β) in (3.169) perform in parallel. Also the behaviors of (θ, β) in (3.170) and (λ, β) in (3.171) are parallel. Refer to Sun (1997) for details. Thus only the model (3.170) is considered here.

The Fisher information matrix of (θ, β) for the Weibull model (3.170) is

$$\mathbf{\Sigma}(\theta, \beta) = \begin{pmatrix} \frac{1}{\theta^2} & -\frac{1+\gamma_1+\log\theta}{\theta\beta} \\ -\frac{1+\gamma_1+\log\theta}{\theta\beta} & \frac{\gamma_2-\gamma_1^2+(1+\gamma_1+\log\theta)^2}{\beta^2} \end{pmatrix}, \tag{3.172}$$

where $\gamma_i = \int_0^\infty [\log(z)]^i e^{-z} dz$. Thus

$$\begin{aligned} |\mathbf{\Sigma}| &= \frac{\gamma_2-\gamma_1^2}{\theta^2\beta^2}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{\gamma_2-\gamma_1^2+(1+\gamma_1+\log\theta)^2}{\beta^2}, \\ |\mathbf{\Sigma}_{22}^c| &= \frac{1}{\theta^2}, \quad |\mathbf{\Sigma}_{12}^c| = |\mathbf{\Sigma}_{21}^c| = -\frac{1+\gamma_1+\log\theta}{\theta\beta}, \end{aligned}$$

and then

$$\begin{aligned} \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{\gamma_2-\gamma_1^2}{\theta^2\{\gamma_2-\gamma_1^2+(1+\gamma_1+\log\theta)^2\}}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} &= \frac{\gamma_2-\gamma_1^2}{\beta^2}, \\ \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = -\frac{\gamma_2-\gamma_1^2}{\theta\beta(1+\gamma_1+\log\theta)}. \end{aligned}$$

Hence the marginal reference priors for θ and β are

$$\pi_1^R(\theta) \propto \frac{g(\theta)}{\theta}, \ \theta > 0, \tag{3.173}$$

$$\pi_2^R(\beta) \propto \frac{1}{\beta}, \ \beta > 0, \tag{3.174}$$

and the independent reference prior for (θ, β) is

$$\pi^R(\theta, \beta) \propto \frac{g(\theta)}{\theta\beta}, \tag{3.175}$$

where

$$g(\theta) = \frac{1}{\sqrt{\gamma_2-\gamma_1^2+(1+\gamma_1+\log\theta)^2}}.$$

It is a first order matching prior for θ, and the reference prior for (θ, β) when θ is the parameter of interest and β is nuisance parameter.

3.13 Weibull Model: Two Independent Samples with the Same Shape Parameter

Consider the stress-strength system, where Y is the strength of a system subject to the stress X. The system is reliable when the applied stress (X) is less than its strength (Y). Thus the reliability of the system is defined as

$$\omega_1 = P(X < Y). \tag{3.176}$$

This is used in many areas, especially in structural and aircraft industries. Sun, Ghosh and Basu (1998) performed the objective Bayesian analysis for ω_1 by using reference and matching priors when both of the stress and strength follow the Weibull distribution. Here we develop the independent reference priors.

Suppose that $X_1, \ldots, X_m$ are *iid* $W(\eta_1, \beta)$ and independently, $Y_1, \ldots, Y_n$ are *iid* $W(\eta_2, \beta)$ with the Weibull density given by (3.164). Then for $x_i > 0, y_j > 0, \eta_k > 0, \beta > 0,\ i = 1, \ldots, m, j = 1, \ldots, n, k = 1, 2$, the joint density of $\boldsymbol{x} = (x_1, \ldots, x_m)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$ is

$$f(\boldsymbol{x}, \boldsymbol{y}|\eta_1, \eta_2, \beta) = \left[\prod_{i=1}^{m} \frac{\beta x_i^{\beta-1}}{\eta_1^{\beta}} \exp\left\{-\left(\frac{x_i}{\eta_1}\right)^{\beta}\right\}\right] \left[\prod_{j=1}^{n} \frac{\beta y_j^{\beta-1}}{\eta_2^{\beta}} \exp\left\{-\left(\frac{y_j}{\eta_2}\right)^{\beta}\right\}\right].$$

First, we derive the independent reference prior for (η_1, η_2, β). The Fisher information matrix of (η_1, η_2, β) is

$$\boldsymbol{\Sigma}(\eta_1, \eta_2, \beta) = \begin{pmatrix} \frac{m\beta^2}{\eta_1^2} & 0 & -\frac{m(1+\gamma_1)}{\eta_1} \\ 0 & \frac{n\beta^2}{\eta_2^2} & -\frac{n(1+\gamma_1)}{\eta_2} \\ -\frac{m(1+\gamma_1)}{\eta_1} & -\frac{n(1+\gamma_1)}{\eta_2} & \frac{(m+n)(\gamma_2+2\gamma_1+1)}{\beta^2} \end{pmatrix}, \tag{3.177}$$

where $\gamma_i = \int_0^{\infty} [\log(z)]^i e^{-z} dz$. Thus

$$|\boldsymbol{\Sigma}| = \frac{mn(m+n)(\gamma_2 - \gamma_1^2)\beta^2}{\eta_1^2 \eta_2^2}, \quad |\boldsymbol{\Sigma}_{11}^c| = \frac{n^2(\gamma_2 - \gamma_1^2) + mn(\gamma_2 + 2\gamma_1 + 1)}{\eta_2^2},$$

$$\begin{aligned}
|\mathbf{\Sigma}_{22}^c| &= \frac{m^2(\gamma_2-\gamma_1^2)+mn(\gamma_2+2\gamma_1+1)}{\eta_1^2}, \quad |\mathbf{\Sigma}_{33}^c| = \frac{mn\beta^4}{\eta_1^2\eta_2^2},\\
|\mathbf{\Sigma}_{12}^c| &= |\mathbf{\Sigma}_{21}^c| = -\frac{mn(1+\gamma_1)^2}{\eta_1\eta_2}, \quad |\mathbf{\Sigma}_{13}^c| = |\mathbf{\Sigma}_{31}^c| = \frac{mn(1+\gamma_1)\beta^2}{\eta_1\eta_2^2},\\
|\mathbf{\Sigma}_{23}^c| &= |\mathbf{\Sigma}_{32}^c| = -\frac{mn(1+\gamma_1)\beta^2}{\eta_1^2\eta_2},
\end{aligned}$$

and then

$$\begin{aligned}
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{1}{\eta_1^2}\left\{\frac{mn(m+n)(\gamma_2-\gamma_1^2)\beta^2}{n^2(\gamma_2-\gamma_1^2)+mn(\gamma_2+2\gamma_1+1)}\right\},\\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} &= \frac{1}{\eta_2^2}\left\{\frac{mn(m+n)(\gamma_2-\gamma_1^2)\beta^2}{m^2(\gamma_2-\gamma_1^2)+mn(\gamma_2+2\gamma_1+1)}\right\},\\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} &= \frac{(m+n)(\gamma_2-\gamma_1^2)}{\beta^2},\\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = -\frac{(m+n)(\gamma_2-\gamma_1^2)\beta^2}{\eta_1\eta_2(1+\gamma_1)^2},\\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{13}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{31}^c|} = \frac{(m+n)(\gamma_2-\gamma_1^2)}{\eta_1(1+\gamma_1)},\\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{23}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{32}^c|} = -\frac{(m+n)(\gamma_2-\gamma_1^2)}{\eta_2(1+\gamma_1)}.
\end{aligned}$$

Hence the marginal reference priors for η_1, η_2 and β are

$$\pi_1^R(\eta_1) \propto \frac{1}{\eta_1}, \ \eta_1 > 0, \tag{3.178}$$

$$\pi_2^R(\eta_2) \propto \frac{1}{\eta_2}, \ \eta_2 > 0, \tag{3.179}$$

$$\pi_3^R(\beta) \propto \frac{1}{\beta}, \ \beta > 0, \tag{3.180}$$

and the independent reference prior for (η_1, η_2, β) is

$$\pi^R(\eta_1, \eta_2, \beta) \propto \frac{1}{\eta_1\eta_2\beta}. \tag{3.181}$$

It is also a first order matching prior for η_1, η_2 and β, and the one-at-a-time reference prior for (η_1, η_2, β) with any ordering.

Consider alternative reparameterization, $\theta_1 = \eta_1^\beta$ and $\theta_2 = \eta_2^\beta$. As you can see in (3.187), the reliability of the system, ω_1 defined as (3.176), is composed of θ_1 and θ_2. The Fisher information matrix of $(\theta_1, \theta_2, \beta)$ is

$$\boldsymbol{\Sigma}(\theta_1, \theta_2, \beta) = \begin{pmatrix} \frac{m}{\theta_1^2} & 0 & -\frac{m(\log\theta_1+1+\gamma_1)}{\theta_1\beta} \\ 0 & \frac{n}{\theta_2^2} & -\frac{n(\log\theta_2+1+\gamma_1)}{\theta_2\beta} \\ -\frac{m(\log\theta_1+1+\gamma_1)}{\theta_1\beta} & -\frac{n(\log\theta_2+1+\gamma_1)}{\theta_2\beta} & K(\theta_1, \theta_2, \beta) \end{pmatrix}, \tag{3.182}$$

where $\gamma_i = \int_0^\infty [\log(z)]^i e^{-z} dz$ and $K(\theta_1, \theta_2, \beta) = \{m(\log\theta_1 + 1 + \gamma_1)^2 + n(\log\theta_2 + 1 + \gamma_1)^2 + (m+n)(\gamma_2 - \gamma_1^2)\}/\beta^2$. Thus

$$\begin{aligned}
|\boldsymbol{\Sigma}| &= \frac{mn(m+n)(\gamma_2 - \gamma_1^2)}{\theta_1^2\theta_2^2\beta^2}, \\
|\boldsymbol{\Sigma}_{11}^c| &= \frac{n\{m(\log\theta_1 + 1 + \gamma_1)^2 + (m+n)(\gamma_2 - \gamma_1^2)\}}{\theta_2^2\beta^2}, \\
|\boldsymbol{\Sigma}_{22}^c| &= \frac{m\{n(\log\theta_2 + 1 + \gamma_1)^2 + (m+n)(\gamma_2 - \gamma_1^2)\}}{\theta_1^2\beta^2}, \\
|\boldsymbol{\Sigma}_{33}^c| &= \frac{mn}{\theta_1^2\theta_2^2}, \\
|\boldsymbol{\Sigma}_{12}^c| &= |\boldsymbol{\Sigma}_{21}^c| = -\frac{mn(\log\theta_1 + 1 + \gamma_1)(\log\theta_2 + 1 + \gamma_1)}{\theta_1\theta_2\beta^2}, \\
|\boldsymbol{\Sigma}_{13}^c| &= |\boldsymbol{\Sigma}_{31}^c| = \frac{mn(\log\theta_1 + 1 + \gamma_1)}{\theta_1\theta_2^2\beta}, \\
|\boldsymbol{\Sigma}_{23}^c| &= |\boldsymbol{\Sigma}_{32}^c| = -\frac{mn(\log\theta_2 + 1 + \gamma_1)}{\theta_1^2\theta_2\beta},
\end{aligned}$$

and then

$$\begin{aligned}
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}^c|} &= \frac{m(m+n)(\gamma_2 - \gamma_1^2)}{\theta_1^2\{m(\log\theta_1 + 1 + \gamma_1)^2 + (m+n)(\gamma_2 - \gamma_1^2)\}}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{22}^c|} &= \frac{n(m+n)(\gamma_2 - \gamma_1^2)}{\theta_2^2\{n(\log\theta_2 + 1 + \gamma_1)^2 + (m+n)(\gamma_2 - \gamma_1^2)\}}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{33}^c|} &= \frac{(m+n)(\gamma_2 - \gamma_1^2)}{\beta^2}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{12}^c|} &= \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{21}^c|} = -\frac{(m+n)(\gamma_2 - \gamma_1^2)}{\theta_1\theta_2(\log\theta_1 + 1 + \gamma_1)(\log\theta_2 + 1 + \gamma_1)}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{13}^c|} &= \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{31}^c|} = \frac{(m+n)(\gamma_2 - \gamma_1^2)}{\beta\theta_1(\log\theta_1 + 1 + \gamma_1)}, \\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{23}^c|} &= \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{32}^c|} = -\frac{(m+n)(\gamma_2 - \gamma_1^2)}{\beta\theta_2(\log\theta_2 + 1 + \gamma_1)}.
\end{aligned}$$

Hence the marginal reference priors for θ_1, θ_2 and β are

$$\pi_1^R(\theta_1) \propto \frac{1}{\theta_1\sqrt{m(\log\theta_1+1+\gamma_1)^2+(m+n)(\gamma_2-\gamma_1^2)}},\ \theta_1>0, \tag{3.183}$$

$$\pi_2^R(\theta_2) \propto \frac{1}{\theta_2\sqrt{n(\log\theta_2+1+\gamma_1)^2+(m+n)(\gamma_2-\gamma_1^2)}},\ \theta_2>0, \tag{3.184}$$

$$\pi_3^R(\beta) \propto \frac{1}{\beta},\ \beta>0, \tag{3.185}$$

and the independent reference prior for $(\theta_1,\theta_2,\beta)$ is

$$\pi^R(\theta_1,\theta_2,\beta)=\pi_1^R(\theta_1)\pi_2^R(\theta_2)\pi_3^R(\beta). \tag{3.186}$$

Next, two independent reference priors under the Weibull stress-strength model are derived by considering two different sets of nuisance parameters. When the stress and strength are Weibull random samples, the parameter of interest in (3.176) can be rewritten as

$$\omega_1=\frac{\eta_1^{-\beta}}{\eta_1^{-\beta}+\eta_2^{-\beta}}=\frac{\eta_2^{\beta}}{\eta_1^{\beta}+\eta_2^{\beta}}. \tag{3.187}$$

Sun, Ghosh and Basu (1998) chose $\omega_2=1/(\eta_1^{-\beta}+\eta_2^{-\beta})=\eta_1^{\beta}\eta_2^{\beta}/(\eta_1^{\beta}+\eta_2^{\beta})$ and β as nuisance parameters and computed various reference priors. The independent reference prior for $(\omega_1,\omega_2,\beta)$ is derived here. The Fisher information matrix of $(\omega_1,\omega_2,\beta)$ was given by Sun, Ghosh and Basu (1998) as follows,

$$\mathbf{\Sigma}(\omega_1,\omega_2,\beta)=(I_{ij})_{3\times3}, \tag{3.188}$$

where

$$\begin{aligned}
I_{11} &= \frac{m}{\omega_1^2}+\frac{n}{(1-\omega_1)^2},\quad I_{12}=\frac{m}{\omega_1\omega_2}-\frac{n}{(1-\omega_1)\omega_2},\\
I_{13} &= \frac{m\{1+\gamma_1-\log(\omega_1\omega_2)\}}{\omega_1\beta}-\frac{n[1+\gamma_1-\log\{(1-\omega_1)\omega_2\}]}{(1-\omega_1)\beta},\\
I_{22} &= \frac{m+n}{\omega_2^2},\quad I_{23}=\frac{(m+n)(1+\gamma_1)-m\log(\omega_1\omega_2)-n\log\{(1-\omega_1)\omega_2\}}{\omega_2\beta},\\
I_{33} &= \frac{(m+n)(\gamma_2-\gamma_1^2)+m\{1+\gamma_1-\log(\omega_1\omega_2)\}^2+n[1+\gamma_1-\log\{(1-\omega_1)\omega_2\}]^2}{\beta^2},
\end{aligned}$$

where $\gamma_i = \int_0^\infty [\log(z)]^i e^{-z} dz$. Thus

$$
\begin{aligned}
|\mathbf{\Sigma}| &= \frac{mn(m+n)(\gamma_2-\gamma_1^2)}{\omega_1^2(1-\omega_1)^2\omega_2^2\beta^2}, \\
|\mathbf{\Sigma}_{11}^c| &= \frac{(m+n)^2(\gamma_2-\gamma_1^2)+mn\log^2\left(\frac{1-\omega_1}{\omega_1}\right)}{\omega_2^2\beta^2}, \\
|\mathbf{\Sigma}_{22}^c| &= \frac{(m+n)(\gamma_2-\gamma_1^2)\{m(1-\omega_1)^2+n\omega_1^2\}}{\omega_1^2(1-\omega_1)^2\beta^2} \\
&\quad + \frac{mn\left[1+\gamma_1+\omega_1\log\left(\frac{1-\omega_1}{\omega_1}\right)-\log\{(1-\omega_1)\omega_2\}\right]^2}{\omega_1^2(1-\omega_1)^2\beta^2}, \\
|\mathbf{\Sigma}_{33}^c| &= \frac{mn}{\omega_1^2(1-\omega_1)^2\omega_2^2},
\end{aligned}
$$

and then

$$
\begin{aligned}
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{an(\gamma_2-\gamma_1^2)}{\omega_1^2(1-\omega_1)^2\left\{\gamma_2-\gamma_1^2+a(1-a)\log^2\left(\frac{1-\omega_1}{\omega_1}\right)\right\}}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} &= \frac{mn(\gamma_2-\gamma_1^2)g(\omega_1,\omega_2)}{\omega_2^2}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} &= \frac{(m+n)(\gamma_2-\gamma_1^2)}{\beta^2},
\end{aligned}
$$

where

$$
g(\omega_1,\omega_2) = \frac{1}{(\gamma_2-\gamma_1^2)\{m(1-\omega_1)^2+n\omega_1^2\}+an\left[1+\gamma_1+\omega_1\log\left(\frac{1-\omega_1}{\omega_1}\right)-\log\{(1-\omega_1)\omega_2\}\right]^2}
$$

and $a = \frac{m}{m+n}$. It is obvious that $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{ii}^c|$, $i = 1, 3$, satisfy Condition (2.3) but $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{22}^c|$ does not. So we cannot apply Theorem 2.1 in direct to this case. Thus we use the iterative algorithm to derive the independent reference prior for $(\omega_1, \omega_2, \beta)$.

Proposition 3.7 *Choose any constants a_j and b_j such that $a_j \to 0$ and $b_j \to 1$ as $j \to \infty$. Then the marginal reference priors for ω_1, ω_2 and β are*

$$
\pi_1^R(\omega_1) \propto \frac{g_1(\omega_1)}{\omega_1(1-\omega_1)},\ \omega_1 \in (0,1), \tag{3.189}
$$

$$
\pi_2^R(\omega_2) = \lim_{j\to\infty}\frac{A_j(\omega_2)}{A_j(1)},\ \omega_2 > 0, \tag{3.190}
$$

$$
\pi_3^R(\beta) \propto \frac{1}{\beta},\ \beta > 0, \tag{3.191}
$$

where

$$\begin{aligned}
A_j(\omega_2) &= \exp\left\{\frac{\int_{a_j}^{b_j} \frac{g_1(\omega_1)}{\omega_1(1-\omega_1)} \log h(\omega_1,\omega_2) d\omega_1}{2\int_{a_j}^{b_j} \frac{g_1(\omega_1)}{\omega_1(1-\omega_1)} d\omega_1}\right\},\\
g_1(\omega_1) &= \frac{1}{\sqrt{\gamma_2-\gamma_1^2+a(1-a)\log^2\left(\frac{1-\omega_1}{\omega_1}\right)}},\\
h(\omega_1,\omega_2) &= \frac{g(\omega_1,\omega_2)}{\omega_2^2},\\
g(\omega_1,\omega_2) &= \frac{1}{(\gamma_2-\gamma_1^2)\{m(1-\omega_1)^2+n\omega_1^2\}+an\left[1+\gamma_1+\omega_1\log\left(\frac{1-\omega_1}{\omega_1}\right)-\log\{(1-\omega_1)\omega_2\}\right]^2}.
\end{aligned}$$

Consequently, the independent reference prior for $(\omega_1, \omega_2, \beta)$ is

$$\pi^R(\omega_1,\omega_2,\beta) = \pi_1^R(\omega_1)\pi_2^R(\omega_2)\pi_3^R(\beta). \tag{3.192}$$

Proof. It is seen that $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{ii}^c|$, $i = 1, 3$, satisfy Condition (2.3) so that (3.189) and (3.191) hold immediately. It is clear that $\pi_1^R(\omega_1)$ and $\pi_3^R(\beta)$ are improper. Thus we need an argument of compact sets and use *Algorithm B′* to derive $\pi_2^R(\omega_2)$ since $|\mathbf{\Sigma}|/|\mathbf{\Sigma}_{22}^c|$ does not meet Condition (2.3). Choose any constants a_j and b_j such that $a_j \to 0$ and $b_j \to 1$ as $j \to \infty$. Also choose any constants c_j and d_j such that $c_j \to 0$ and $d_j \to \infty$ as $j \to \infty$. Then by applying *Algorithm B′*, it can be shown that

$$\begin{aligned}
\pi_{2j}(\omega_2) &\propto \exp\left\{\frac{\int_{c_j}^{d_j}\int_{a_j}^{b_j} \pi_1^R(\omega_1)\pi_3^R(\beta)\log\left(\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|}\right) d\omega_1 d\beta}{2\int_{c_j}^{d_j}\int_{a_j}^{b_j} \pi_1^R(\omega_1)\pi_3^R(\beta) d\omega_1 d\beta}\right\}\\
&= \exp\left\{\frac{\int_{c_j}^{d_j}\int_{a_j}^{b_j} \frac{g_1(\omega_1)}{\beta\omega_1(1-\omega_1)}\log h(\omega_1,\omega_2) d\omega_1 d\beta}{2\int_{c_j}^{d_j}\int_{a_j}^{b_j} \frac{g_1(\omega_1)}{\beta\omega_1(1-\omega_1)} d\omega_1 d\beta}\right\}\\
&= \exp\left\{\frac{\int_{a_j}^{b_j} \frac{g_1(\omega_1)}{\omega_1(1-\omega_1)}\log h(\omega_1,\omega_2) d\omega_1}{2\int_{a_j}^{b_j} \frac{g_1(\omega_1)}{\omega_1(1-\omega_1)} d\omega_1}\right\}\\
&\equiv A_j(\omega_2).
\end{aligned}$$

Set $\omega_2^0 = 1$. Thus

$$\pi_2^R(\omega_2) = \lim_{j\to\infty}\frac{\pi_{2j}(\omega_2)}{\pi_{2j}(\omega_2^0)} = \lim_{j\to\infty}\frac{\pi_{2j}(\omega_2)}{\pi_{2j}(1)} = \lim_{j\to\infty}\frac{A_j(\omega_2)}{A_j(1)}.$$

The result then holds. $\square$

The independent reference prior for $(\omega_1, \omega_2, \beta)$ does not have a closed form because it is practically impossible to calculate $A_j(\cdot)$ in $\pi_2^R(\omega_2)$. Thus we now choose a new set of nuisance parameters so that the independent reference prior using it has a closed form.

Consider $\omega_2 = \eta_2^\beta$ and β as nuisance parameters. Define $\gamma_i = \int_0^\infty [\log(z)]^i e^{-z} dz$ for $i \geq 0$. Then the Fisher information matrix of $(\omega_1, \omega_2, \beta)$ is

$$\mathbf{\Sigma}(\omega_1, \omega_2, \beta) = (I_{ij})_{3\times 3}, \tag{3.193}$$

where

$$\begin{aligned}
I_{11} &= \frac{m}{\omega_1^2(1-\omega_1)^2}, \quad I_{12} = -\frac{m}{\omega_1(1-\omega_1)\omega_2}, \\
I_{13} &= \frac{m\left[1+\gamma_1+\log\left\{\frac{(1-\omega_1)\omega_2}{\omega_1}\right\}\right]}{\omega_1(1-\omega_1)\beta}, \quad I_{22} = \frac{m+n}{\omega_2^2}, \\
I_{23} &= -\frac{m\left[1+\gamma_1+\log\left\{\frac{(1-\omega_1)\omega_2}{\omega_1}\right\}\right] + n(1+\gamma_1+\log\omega_2)}{\omega_2\beta}, \\
I_{33} &= \frac{(m+n)(\gamma_2-\gamma_1^2) + m\left[1+\gamma_1+\log\left\{\frac{(1-\omega_1)\omega_2}{\omega_1}\right\}\right]^2 + n(1+\gamma_1+\log\omega_2)^2}{\beta^2}.
\end{aligned}$$

Thus

$$\begin{aligned}
|\mathbf{\Sigma}| &= \frac{mn(m+n)(\gamma_2-\gamma_1^2)}{\omega_1^2(1-\omega_1)^2\omega_2^2\beta^2}, \\
|\mathbf{\Sigma}_{11}^c| &= \frac{(m+n)^2(\gamma_2-\gamma_1^2) + mn\log^2\left(\frac{1-\omega_1}{\omega_1}\right)}{\omega_2^2\beta^2}, \\
|\mathbf{\Sigma}_{22}^c| &= \frac{m(m+n)(\gamma_2-\gamma_1^2) + mn(1+\gamma_1+\log\omega_2)^2}{\omega_1^2(1-\omega_1)^2\beta^2}, \\
|\mathbf{\Sigma}_{33}^c| &= \frac{mn}{\omega_1^2(1-\omega_1)^2\omega_2^2}, \\
|\mathbf{\Sigma}_{12}^c| &= |\mathbf{\Sigma}_{21}^c| = -\frac{m(m+n)(\gamma_2-\gamma_1^2) - mn(1+\gamma_1+\log\omega_2)\log\left(\frac{1-\omega_1}{\omega_1}\right)}{\omega_1(1-\omega_1)\omega_2\beta^2}, \\
|\mathbf{\Sigma}_{13}^c| &= |\mathbf{\Sigma}_{31}^c| = \frac{mn\log\left(\frac{\omega_1}{1-\omega_1}\right)}{\omega_1(1-\omega_1)\omega_2^2\beta},
\end{aligned}$$

$$|\mathbf{\Sigma}_{23}^c| = |\mathbf{\Sigma}_{32}^c| = -\frac{mn(1+\gamma_1+\log\omega_2)}{\omega_1^2(1-\omega_1)^2\omega_2\beta}.$$

Let $a = m/(m+n)$. Then we have

$$\begin{aligned}
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{an(\gamma_2-\gamma_1^2)}{\omega_1^2(1-\omega_1)^2\left\{\gamma_2-\gamma_1^2+a(1-a)\log^2\left(\frac{1-\omega_1}{\omega_1}\right)\right\}}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} &= \frac{n(\gamma_2-\gamma_1^2)}{\omega_2^2\left\{\gamma_2-\gamma_1^2+(1-a)(1+\gamma_1+\log\omega_2)^2\right\}}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} &= \frac{(m+n)(\gamma_2-\gamma_1^2)}{\beta^2}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = -\frac{n(\gamma_2-\gamma_1^2)}{\omega_2\omega_1(1-\omega_1)\left\{(\gamma_2-\gamma_1^2)-(1-a)(1+\gamma_1+\log\omega_2)\log\left(\frac{1-\omega_1}{\omega_1}\right)\right\}}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{13}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{31}^c|} = \frac{(m+n)(\gamma_2-\gamma_1^2)}{\beta\omega_1(1-\omega_1)\log\left(\frac{\omega_1}{1-\omega_1}\right)}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{23}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{32}^c|} = -\frac{(m+n)(\gamma_2-\gamma_1^2)}{\beta\omega_2(1+\gamma_1+\log\omega_2)}.
\end{aligned}$$

Thus the marginal reference priors for ω_1, ω_2 and β are

$$\pi_1^R(\omega_1) \propto \frac{g_1(\omega_1)}{\omega_1(1-\omega_1)},\ \omega_1 \in (0,1), \tag{3.194}$$

$$\pi_2^R(\omega_2) \propto \frac{g_2(\omega_2)}{\omega_2},\ \omega_2 > 0, \tag{3.195}$$

$$\pi_3^R(\beta) \propto \frac{1}{\beta},\ \beta > 0, \tag{3.196}$$

where

$$\begin{aligned}
g_1(\omega_1) &= \frac{1}{\sqrt{\gamma_2-\gamma_1^2+a(1-a)\log^2\left(\frac{1-\omega_1}{\omega_1}\right)}}, \\
g_2(\omega_2) &= \frac{1}{\sqrt{\gamma_2-\gamma_1^2+(1-a)(1+\gamma_1+\log\omega_2)^2}}.
\end{aligned}$$

Consequently, the independent reference prior for $(\omega_1, \omega_2, \beta)$ is

$$\pi^R(\omega_1,\omega_2,\beta) \propto \frac{g_1(\omega_1)g_2(\omega_2)}{(1-\omega_1)\omega_1\omega_2\beta}. \tag{3.197}$$

The reference priors for $(\omega_1, \omega_2, \beta)$ when $\omega_2 = \eta_2^\beta$ are the same as those when $\omega_2 = 1/(\eta_1^{-\beta} + \eta_2^{-\beta}) = \eta_1^\beta\eta_2^\beta/(\eta_1^\beta+\eta_2^\beta)$. However the independent reference prior for $(\omega_1, \omega_2, \beta)$, given by (3.197), is different from all of the reference priors.

3.14 One-way Random Effects ANOVA Model

The unbalanced one-way random effects ANOVA model is

$$X_{ij} = \mu + \alpha_i + \epsilon_{ij}, \quad i = 1, \ldots, k, j = 1, \ldots, n_i,$$

where $\mu \in \mathbb{R}$, α_i *iid* $N(0, \tau^2), \tau > 0$, ϵ_{ij} *iid* $N(0, \sigma^2), \sigma > 0$, and α_i and ϵ_{ij} are mutually independent.

Let $\eta = \frac{\tau^2}{\sigma^2}$. Then the Fisher information matrix of (η, σ^2, μ) is

$$\mathbf{\Sigma}(\eta, \sigma^2, \mu) = \begin{pmatrix} \frac{s_{22}(\eta)}{2} & \frac{s_{11}(\eta)}{2\sigma^2} & 0 \\ \frac{s_{11}(\eta)}{2\sigma^2} & \frac{n}{2\sigma^4} & 0 \\ 0 & 0 & \frac{s_{11}(\eta)}{\sigma^2} \end{pmatrix}, \tag{3.198}$$

where $n = \sum_{i=1}^{k} n_i$ and $s_{pq}(x) = \sum_{i=1}^{k} \frac{n_i^p}{(1+n_i x)^q}$. Thus

$$\begin{aligned}
|\mathbf{\Sigma}| &= \frac{s_{11}(\eta)\{n s_{22}(\eta) - s_{11}(\eta)^2\}}{4\sigma^6}, \quad |\mathbf{\Sigma}_{11}^c| = \frac{n s_{11}(\eta)}{2\sigma^6}, \\
|\mathbf{\Sigma}_{22}^c| &= \frac{s_{11}(\eta) s_{22}(\eta)}{2\sigma^2}, \quad |\mathbf{\Sigma}_{33}^c| = \frac{n s_{22}(\eta) - s_{11}(\eta)^2}{4\sigma^4}, \\
|\mathbf{\Sigma}_{ij}^c| &= \begin{cases} \frac{s_{11}(\eta)^2}{2\sigma^4}, & \text{if } (i,j) = (1,2), (2,1), \\ 0, & \text{otherwise,} \end{cases}
\end{aligned}$$

and then

$$\begin{aligned}
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{11}^c|} &= \frac{n s_{22}(\eta) - s_{11}(\eta)^2}{2n}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{22}^c|} &= \frac{1}{\sigma^4} \left\{ \frac{n s_{22}(\eta) - s_{11}(\eta)^2}{2 s_{22}(\eta)} \right\}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{33}^c|} &= \frac{s_{11}(\eta)}{\sigma^2}, \\
\frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{12}^c|} &= \frac{|\mathbf{\Sigma}|}{|\mathbf{\Sigma}_{21}^c|} = \frac{n s_{22}(\eta) - s_{11}(\eta)^2}{2\sigma^2 s_{11}(\eta)}.
\end{aligned}$$

Hence the marginal reference priors for η, σ^2 and μ are

$$\pi_1^R(\eta) \propto \sqrt{n s_{22}(\eta) - s_{11}(\eta)^2}, \; \eta > 0, \tag{3.199}$$

$$\pi_2^R(\sigma^2) \propto \frac{1}{\sigma^2},\ \sigma^2 > 0, \tag{3.200}$$

$$\pi_3^R(\mu) \propto 1,\ \mu \in \mathbb{R}, \tag{3.201}$$

and the independent reference prior for (η, σ^2, μ) is

$$\pi^R(\eta, \sigma^2, \mu) \propto \frac{\sqrt{ns_{22}(\eta) - s_{11}(\eta)^2}}{\sigma^2}. \tag{3.202}$$

It is also a first order matching prior for η and μ, and the one-at-a-time reference prior for (η, σ^2, μ) in the ordering of $\{\mu, \eta, \sigma^2\}$, $\{\eta, \mu, \sigma^2\}$ and $\{\eta, \sigma^2, \mu\}$.

3.15 Two-parameter Exponential Family

Referring to Sun and Ye (1996), a two-parameter exponential family has a density of

$$f(x|\mu, \beta) = \exp\left\{\beta[U_1(x) + g(\mu)U_2(x) + \mu G_1'(\mu) - G_1(\mu)] - G_2(\beta)\right\},$$

where both $G_1(\mu) = -\int g(\mu)d\mu$ and $G_2(\beta), \beta < 0$ are infinitely differentiable and strictly convex functions. See Sun and Ye (1996) for details on the two-parameter exponential family. The gamma [Section 3.5], inverse Gaussian [Section 3.6], lognormal [Section 3.7] and normal [Section 3.8] models are some well-known examples of the two-parameter exponential family. The Fisher information matrix of (μ, β) is

$$\mathbf{\Sigma}(\mu, \beta) = \begin{pmatrix} -\beta G_1''(\mu) & 0 \\ 0 & G_2''(\beta) \end{pmatrix}. \tag{3.203}$$

Hence the marginal reference priors for μ and β are

$$\pi_1^R(\mu) \propto \sqrt{G_1''(\mu)}, \tag{3.204}$$

$$\pi_2^R(\beta) \propto \sqrt{G_2''(\beta)}, \tag{3.205}$$

and the independent reference prior for (μ, β) is

$$\pi^R(\mu, \beta) \propto \sqrt{G_1''(\mu) G_2''(\beta)}. \tag{3.206}$$

It is a first order matching prior for μ and β. The independent reference prior for (μ, β) is the same as the reference prior for (μ, β), derived by Sun and Ye (1996), when one of the parameters μ or β is the interest parameter and the other is nuisance parameter.

3.16 Proper Two-parameter Dispersion Model

A two-parameter dispersion model is defined as

$$f(y|\mu, \lambda) = c(\lambda, y) \exp\{\lambda t(y, \mu)\}$$

for some functions $c(\cdot)$ (> 0) and $t(\cdot)$. When $c(\lambda, y)$ can be expressed as $a(\lambda)b(y)$, such models are called as proper dispersion models. In general, μ is the mean of the distribution. A two-parameter exponential family in Section 3.15 is an example of proper two-parameter dispersion model. Garvan and Ghosh (1997) derived noninformative priors, such as Jeffreys, reference and probability matching priors, for two-parameter dispersion models. In this section, only a proper two-parameter dispersion model is considered. Then the Fisher information matrix of (μ, λ) has the form of

$$\mathbf{\Sigma}(\mu, \lambda) = \begin{pmatrix} h_{11}(\mu) h_{12}(\lambda) & 0 \\ 0 & h_{21}(\lambda) h_{22}(\mu) \end{pmatrix} \tag{3.207}$$

for some positive functions $h_{11}(\cdot)$, $h_{12}(\cdot)$, $h_{21}(\cdot)$ and $h_{22}(\cdot)$. Thus the marginal reference priors for μ and λ are

$$\pi_1^R(\mu) \propto \sqrt{h_{11}(\mu)}, \tag{3.208}$$

$$\pi_2^R(\lambda) \propto \sqrt{h_{21}(\lambda)}, \tag{3.209}$$

and the independent reference prior for (μ, λ) is

$$\pi^R(\mu, \lambda) \propto \sqrt{h_{11}(\mu) h_{21}(\lambda)}. \tag{3.210}$$

It is also a first order matching prior for μ and λ. The independent reference prior for (μ, λ) is the same as the reference prior, which is computed by Garvan and Ghosh (1997), whichever of μ and λ is the parameter of interest.

3.16.1 Typical Examples

We provide the summary of component functions for some typical members of the proper dispersion family by referring to Garvan and Ghosh (1997). The independent reference priors are then derived.

For the Fisher-von Mises distribution,

$$\begin{aligned} a(\lambda) &= \frac{1}{2\pi I_0(\lambda)}, \quad b(y) = 1, \quad t(y, \mu) = \cos(y - \mu), \\ h_{11}(\mu) &= 1, \quad h_{21}(\lambda) = 1 - \frac{A(\lambda)}{\lambda} - A^2(\lambda), \end{aligned}$$

where $I_\nu(\lambda) = (1/\pi) \int_0^\pi e^{\lambda \cos(x)} \cos(\nu x) dx$ and $A(\lambda) = I_1(\lambda)/I_0(\lambda)$. Thus the marginal reference priors for μ and λ are

$$\pi_1^R(\mu) \propto 1, \tag{3.211}$$

$$\pi_2^R(\lambda) \propto \sqrt{1 - \frac{A(\lambda)}{\lambda} - A^2(\lambda)}, \tag{3.212}$$

and the independent reference prior for (μ, λ) is

$$\pi^R(\mu, \lambda) \propto \sqrt{1 - \frac{A(\lambda)}{\lambda} - A^2(\lambda)}. \tag{3.213}$$

For the Student-t distribution,

$$a(\lambda) = \frac{\Gamma(\lambda)}{\sqrt{\pi}\, \Gamma(\lambda - \frac{1}{2})}, \quad b(y) = 1, \quad t(y, \mu) = -\log\{1 + (y - \mu)^2\},$$

$$h_{11}(\mu) = 1, \quad h_{21}(\lambda) = \frac{d^2}{d\lambda^2} \log \frac{\Gamma(\lambda)}{\Gamma(\lambda - \frac{1}{2})}.$$

Thus the marginal reference priors for μ and λ are

$$\pi_1^R(\mu) \propto 1, \tag{3.214}$$

$$\pi_2^R(\lambda) \propto \sqrt{\frac{d^2}{d\lambda^2} \log \frac{\Gamma(\lambda)}{\Gamma(\lambda - \frac{1}{2})}}, \tag{3.215}$$

and the independent reference prior for (μ, λ) is

$$\pi^R(\mu, \lambda) \propto \sqrt{\frac{d^2}{d\lambda^2} \log \frac{\Gamma(\lambda)}{\Gamma(\lambda - \frac{1}{2})}}. \tag{3.216}$$

For the power family,

$$a(\lambda) = \frac{\lambda^\gamma}{2\gamma\Gamma(\lambda)}, \quad b(y) = 1, \quad t(y, \mu) = -|y - \mu|^\delta,$$
$$h_{11}(\mu) = 1, \quad h_{21}(\lambda) = \frac{1}{\lambda^2}.$$

Thus the marginal reference priors for μ and λ are

$$\pi_1^R(\mu) \propto 1, \tag{3.217}$$

$$\pi_2^R(\lambda) \propto \frac{1}{\lambda}, \tag{3.218}$$

and the independent reference prior for (μ, λ) is

$$\pi^R(\mu, \lambda) \propto \frac{1}{\lambda}. \tag{3.219}$$

For the McCullagh distribution,

$$a(\lambda) = \frac{1}{B(\lambda + \frac{1}{2}, \frac{1}{2})}, \quad b(y) = \frac{1}{\sqrt{1 - y^2}}, \quad t(y, \mu) = \log \frac{1 - y^2}{1 - 2y\mu + \mu^2},$$
$$h_{11}(\mu) = \frac{1}{1 - \mu^2}, \quad h_{21}(\lambda) = \frac{d^2}{d\lambda^2} \log B\left(\lambda + \frac{1}{2}, \frac{1}{2}\right),$$

where $B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} = \int_0^1 u^{p-1}(1-u)^{q-1}du$. Thus the marginal reference priors for μ and λ are

$$\pi_1^R(\mu) \propto \frac{1}{\sqrt{1-\mu^2}}, \tag{3.220}$$

$$\pi_2^R(\lambda) \propto \sqrt{\frac{d^2}{d\lambda^2}\log B\left(\lambda+\frac{1}{2},\frac{1}{2}\right)}, \tag{3.221}$$

and the independent reference prior for (μ, λ) is

$$\pi^R(\mu,\lambda) \propto \sqrt{\frac{\frac{d^2}{d\lambda^2}\log B(\lambda+\frac{1}{2},\frac{1}{2})}{1-\mu^2}}. \tag{3.222}$$

3.17 Student-t Regression Model

The Student-t regression model is given by, for $\boldsymbol{x}_i \in \mathbb{R}^p$ and $\boldsymbol{\beta} \in \mathbb{R}^p$,

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \epsilon_i, \quad i = 1, \dots, n,$$

where ϵ_i *iid* $t_\nu(0, \sigma^2)$, $\nu > 0$, $\sigma > 0$. Ferreira (2007) derived the independence Jeffreys prior and Jeffreys-rule prior for $(\boldsymbol{\beta}, \sigma, \nu)$. He also computed the one-at-a-time reference priors for $(\boldsymbol{\beta}, \sigma, \nu)$ by considering the different orders of the parameters. The Fisher information matrix of $(\boldsymbol{\beta}, \sigma, \nu)$ is given by

$$\boldsymbol{\Sigma}(\boldsymbol{\beta},\sigma,\nu) = \begin{pmatrix} \frac{1}{\sigma^2}\frac{\nu+1}{\nu+3}\sum_{i=1}^n \boldsymbol{x}_i\boldsymbol{x}_i' & 0 & 0 \\ 0 & \frac{2n}{\sigma^2}\frac{\nu}{\nu+3} & -\frac{2n}{\sigma}\frac{1}{(\nu+1)(\nu+3)} \\ 0 & -\frac{2n}{\sigma}\frac{1}{(\nu+1)(\nu+3)} & \frac{n}{4}\phi(\nu) \end{pmatrix}, \tag{3.223}$$

where

$$\phi(\nu) = \psi'\left(\frac{\nu}{2}\right) - \psi'\left(\frac{\nu+1}{2}\right) - \frac{2(\nu+5)}{\nu(\nu+1)(\nu+3)}$$

with $\psi(a) = \frac{d}{da}\log\Gamma(a)$ and $\psi'(a) = \frac{d}{da}\psi(a)$. Thus

$$\begin{aligned}
|\boldsymbol{\Sigma}| &= \frac{4n^2}{\sigma^2}\frac{1}{\nu+3}\left|\frac{1}{\sigma^2}\frac{\nu+1}{\nu+3}\sum_{i=1}^n \boldsymbol{x}_i\boldsymbol{x}_i'\right|\left\{\frac{\nu}{8}\phi(\nu) - \frac{1}{(\nu+1)^2(\nu+3)}\right\},\\
|\boldsymbol{\Sigma}_{11}^c| &= \frac{4n^2}{\sigma^2}\frac{1}{\nu+3}\left\{\frac{\nu}{8}\phi(\nu) - \frac{1}{(\nu+1)^2(\nu+3)}\right\},\\
|\boldsymbol{\Sigma}_{22}^c| &= \frac{n}{4}\phi(\nu)\left|\frac{1}{\sigma^2}\frac{\nu+1}{\nu+3}\sum_{i=1}^n \boldsymbol{x}_i\boldsymbol{x}_i'\right|,\\
|\boldsymbol{\Sigma}_{33}^c| &= \frac{2n}{\sigma^2}\frac{\nu}{\nu+3}\left|\frac{1}{\sigma^2}\frac{\nu+1}{\nu+3}\sum_{i=1}^n \boldsymbol{x}_i\boldsymbol{x}_i'\right|,\\
|\boldsymbol{\Sigma}_{ij}^c| &= \begin{cases} -\frac{2n}{\sigma}\frac{1}{(\nu+1)(\nu+3)}\left|\frac{1}{\sigma^2}\frac{\nu+1}{\nu+3}\sum_{i=1}^n \boldsymbol{x}_i\boldsymbol{x}_i'\right|, & \text{if } (i,j) = (2,3),(3,2),\\ 0, & \text{otherwise,}\end{cases}
\end{aligned}$$

and then

$$\begin{aligned}
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{11}^c|} &= \left|\frac{1}{\sigma^2}\frac{\nu+1}{\nu+3}\sum_{i=1}^n \boldsymbol{x}_i\boldsymbol{x}_i'\right|,\\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{22}^c|} &= \frac{4n}{\sigma^2}\left\{\frac{\nu}{8(\nu+3)} - \frac{1}{(\nu+1)^2(\nu+3)^2\phi(\nu)}\right\},\\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{33}^c|} &= \frac{n}{2}\left\{\frac{\phi(\nu)}{8} - \frac{1}{\nu(\nu+1)^2(\nu+3)}\right\} = \frac{n}{16}\left\{\psi'\left(\frac{\nu}{2}\right) - \psi'\left(\frac{\nu+1}{2}\right) - \frac{2(\nu+3)}{\nu(\nu+1)^2}\right\},\\
\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{23}^c|} &= \frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}_{32}^c|} = -\frac{2n}{\sigma}\left\{\frac{\nu(\nu+1)}{8}\phi(\nu) - \frac{1}{(\nu+1)(\nu+3)}\right\}.
\end{aligned}$$

Hence the marginal reference priors for $\boldsymbol{\beta}$, σ and ν are

$$\pi_1^R(\boldsymbol{\beta}) \propto 1,\ \boldsymbol{\beta} \in \mathbb{R}^p, \tag{3.224}$$

$$\pi_2^R(\sigma) \propto \frac{1}{\sigma},\ \sigma > 0, \tag{3.225}$$

$$\pi_3^R(\nu) \propto \sqrt{\psi'\left(\frac{\nu}{2}\right) - \psi'\left(\frac{\nu+1}{2}\right) - \frac{2(\nu+3)}{\nu(\nu+1)^2}},\ \nu > 0. \tag{3.226}$$

Consequently, the independent reference prior for $(\boldsymbol{\beta}, \sigma, \nu)$ is

$$\pi^R(\boldsymbol{\beta}, \sigma, \nu) \propto \frac{1}{\sigma}\sqrt{\psi'\left(\frac{\nu}{2}\right) - \psi'\left(\frac{\nu+1}{2}\right) - \frac{2(\nu+3)}{\nu(\nu+1)^2}}. \tag{3.227}$$

By Ferreira (2007), it is the same as the one-at-a-time reference prior for $(\boldsymbol{\beta}, \sigma, \nu)$ in the ordering of $\{\boldsymbol{\beta}, \nu, \sigma\}$, $\{\nu, \boldsymbol{\beta}, \sigma\}$ and $\{\nu, \sigma, \boldsymbol{\beta}\}$. Assuming that $\boldsymbol{x}_i$ and $\boldsymbol{\beta}$ are real scalars such that $x_i \in \mathbb{R}$ and $\beta \in \mathbb{R}$, the independent reference prior for $(\boldsymbol{\beta}, \sigma, \nu)$ is also a first order matching prior for β and ν.

3.18 Zero-inflated Poisson (ZIP) Model

A zero-inflated Poisson (ZIP) distribution is a useful model for count data which include more zero counts than compatible with the Poisson model. Accoding to Bayarri, Berger and Datta (2007), a ZIP mass function is given by

$$f(x \mid \lambda, p) = pI(x = 0) + (1 - p) f_0(x|\lambda), \quad x = 0, 1, 2, \ldots, \tag{3.228}$$

where $0 < p \leq 1$, $\lambda > 0$, $I(\cdot)$ is the indicator function and

$$f_0(x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots,$$

is the Poisson probability density function. The parameter p is called the *zero-inflated parameter*. If $p = 0$, the ZIP density function is the same as the Poisson density function.

Bayarri, Berger and Datta (2007) conducted the objective testing of a regular Poisson versus a ZIP model using objective Bayesian methodology. They used

$$\mathcal{H}_0 : p = 0 \text{ versus } \mathcal{H}_1 : p > 0$$

as the null and alternative hypotheses, respectively. They derived two objective priors for (λ, p) and computed the Bayes factor of a ZIP to a Poisson model by using the objective priors.

Here we consider a new ZIP model in which the parameters are orthogonal so that an independent reference prior should have a closed form. By Bayarri, Berger and Datta (2007), we can rewrite $f(x \mid \lambda, p)$, given by (3.228), as

$$f^*(x \mid \lambda, p^*) = p^* I(x=0) + (1-p^*) f^T(x|\lambda), \quad x = 0, 1, 2, \ldots, \tag{3.229}$$

where $p^* = p + (1-p)e^{-\lambda}$, $e^{-\lambda} < p^* \leq 1$, and

$$f^T(x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!(1-e^{-\lambda})}, \quad x = 1, 2, \ldots,$$

is the zero-truncated version of the Poisson density function with parameter λ. The parameter of λ in the new ZIP model has the same meaning as that in the original ZIP model. Clearly, if $p^* = e^{-\lambda}$, then the new ZIP density function is equivalent to the Poisson density function. Bayarri, Berger and Datta (2007) derived two objective priors for (λ, p^*) which are given by

$$\pi^l(\lambda, p^*) \propto \frac{k(\lambda)^l}{\sqrt{\lambda}} \frac{I(e^{-\lambda} < p^* \leq 1)}{1-e^{-\lambda}}, \quad l = 0 \text{ or } 1, \tag{3.230}$$

where

$$k(\lambda) = \frac{\sqrt{1-(\lambda+1)e^{-\lambda}}}{1-e^{-\lambda}}.$$

To derive the independent reference prior for (λ, p^*), we see the Fisher information matrix of (λ, p^*) given by

$$\mathbf{\Sigma}(\lambda, p^*) = \begin{pmatrix} \frac{\{1-(\lambda+1)e^{-\lambda}\}(1-p^*)}{\lambda(1-e^{-\lambda})^2} & 0 \\ 0 & \frac{1}{p^*(1-p^*)} \end{pmatrix}. \tag{3.231}$$

It is clear that λ and p^* are orthogonal since the Fisher information matrix of (λ, p^*) is diagonal. The marginal reference priors for λ and p^* are then

$$\pi_1^R(\lambda) \propto \frac{\sqrt{1-(\lambda+1)e^{-\lambda}}}{(1-e^{-\lambda})\sqrt{\lambda}}, \ \lambda > 0, \tag{3.232}$$

$$\pi_2^R(p^*) \propto \frac{1}{\sqrt{p^*(1-p^*)}}, \ e^{-\lambda} < p^* \leq 1. \tag{3.233}$$

Consequently, the independent reference prior for (λ, p^*) is

$$\pi^R(\lambda, p^*) \propto \frac{k(\lambda)}{\sqrt{\lambda}} \frac{I(e^{-\lambda} < p^* \leq 1)}{\sqrt{p^*(1-p^*)}}, \tag{3.234}$$

where

$$k(\lambda) = \frac{\sqrt{1-(\lambda+1)e^{-\lambda}}}{1-e^{-\lambda}}.$$

It is a first order matching prior for λ and p^*. It is seen that the independent reference prior for (λ, p^*) is different from the two objective priors, given by (3.230), derived by Bayarri, Berger and Datta (2007).

Chapter 4

Non-regular Cases

4.1 Setup

In Chapter 2 and 3, we provided the results for regular cases, where the data has common support and the Fisher information is available. In this chapter, we consider some types of non-regular cases, where the support of the data is monotonically decreasing or monotonically increasing in a non-regular type parameter, and the Fisher information matrix cannot be computed. An iterative algorithm to derive an independent reference prior for some types of non-regular cases is presented. We then propose a sufficient condition under which we obtain a closed form of the independent reference prior without going through the iterations. A sufficient condition under which the independent reference prior agrees to a first order matching prior is also given. We derive the independent reference priors in some examples in Section 4.2.

Ghosal and Samanta (1997) derived the reference prior for a one-parameter family of discontinuous densities where the support of the data is monotonically decreasing or monotonically increasing in the parameter. Ghosal (1997) also computed the reference priors in multi-parameter non-regular cases where some regular type parameters are added to the one-parameter family of discontinuous densities used by Ghosal and Samanta (1997).

Now we propose an iterative algorithm to find an independent reference prior in non-regular cases considered by Ghosal (1997). It is an extension of *Algorithm A* introduced by Sun and Berger (1998) for non-regular cases. We consider a multi-parameter family of discontinuous densities used in Ghosal (1997). Suppose a density $f(x;\theta,\boldsymbol{\phi})$, where $\theta \in \Theta \subset \mathbb{R}, \boldsymbol{\phi} = (\varphi_1,\ldots,\varphi_d)' \in \boldsymbol{\Phi} \subset \mathbb{R}^d$. The family $\{f(x;\theta,\boldsymbol{\phi}) : \theta \in \Theta, \boldsymbol{\phi} \in \boldsymbol{\Phi}\}$ is regular with respect to $\boldsymbol{\phi}$ and non-regular with respect to θ where the support of X is either monotonically decreasing or monotonically increasing in θ. In other words, $x \mapsto f(x;\theta,\boldsymbol{\phi})$ is discontinuous at some points which depend on θ only, while for fixed θ, the family $\{f(x;\theta,\boldsymbol{\phi}) : \boldsymbol{\phi} \in \boldsymbol{\Phi}\}$ is regular.

Define

$$c(\theta,\boldsymbol{\phi}) = E\left\{\frac{\partial}{\partial\theta}\log f(X;\theta,\boldsymbol{\phi})\right\}, \tag{4.1}$$

$$\boldsymbol{\lambda}(\theta,\boldsymbol{\phi}) = -E\left\{\frac{\partial^2}{\partial\varphi_j\partial\varphi_k}\log f(X;\theta,\boldsymbol{\phi})\right\}, \quad j,k = 1,\ldots,d, \tag{4.2}$$

where E is the expectation over X given θ and $\boldsymbol{\phi}$.

Algorithm C:

Step 0. Choose any initial nonzero marginal prior density for $\boldsymbol{\phi}$, namely $\pi_2^{(0)}(\boldsymbol{\phi})$.

Step 1. Define an interim prior density for θ by

$$\pi_1^{(1)}(\theta) \propto \exp\left\{\int \pi_2^{(0)}(\boldsymbol{\phi})\log c(\theta,\boldsymbol{\phi})d\boldsymbol{\phi}\right\}.$$

Step 2. Define an interim prior density for $\boldsymbol{\phi}$ by

$$\pi_2^{(1)}(\boldsymbol{\phi}) \propto \exp\left\{\frac{1}{2}\int \pi_1^{(1)}(\theta)\log|\boldsymbol{\lambda}(\theta,\boldsymbol{\phi})|d\theta\right\}.$$

Replace $\pi_2^{(0)}$ in *Step 0* by $\pi_2^{(1)}$ and repeat *Step 1* and *2* to obtain $\pi_i^{(2)}$ for $i = 1, 2$. Finally, the sequences $\{\pi_i^{(l)}\}_{l\geq 1}$, $i = 1, 2$, are generated. The desired marginal reference priors will be the limits

$$\pi_i^R = \lim_{l\to\infty}\pi_i^{(l)}, \quad i = 1, 2,$$

if the limits exist. When we apply the iterative algorithm to derive an independent reference prior, the iterative algorithm might need to operate on compact sets as the regular cases.

The following theorem provides a sufficient condition under which we can have a closed form of the independent reference prior for non-regular cases without doing the iterations.

Theorem 4.1 *Suppose*

$$\begin{aligned} c(\theta,\boldsymbol{\phi}) &= c_1(\theta)c_2(\boldsymbol{\phi}), \\ |\boldsymbol{\lambda}(\theta,\boldsymbol{\phi})| &= \lambda_1(\theta)\lambda_2(\boldsymbol{\phi}), \end{aligned} \tag{4.3}$$

where

$$\begin{aligned} c(\theta,\boldsymbol{\phi}) &= E\left\{\frac{\partial}{\partial\theta}\log f(X;\theta,\boldsymbol{\phi})\right\}, \\ \boldsymbol{\lambda}(\theta,\boldsymbol{\phi}) &= -E\left\{\frac{\partial^2}{\partial\varphi_j\partial\varphi_k}\log f(X;\theta,\boldsymbol{\phi})\right\}, \quad j,k = 1,\ldots,d. \end{aligned}$$

Then the independent reference prior for $(\theta,\boldsymbol{\phi})$ is

$$\pi^R(\theta,\boldsymbol{\phi}) = \pi_1^R(\theta)\pi_2^R(\boldsymbol{\phi}), \tag{4.4}$$

where the marginal reference priors for θ and $\boldsymbol{\phi}$ are

$$\pi_1^R(\theta) \propto c_1(\theta), \tag{4.5}$$

$$\pi_2^R(\boldsymbol{\phi}) \propto \sqrt{\lambda_2(\boldsymbol{\phi})}. \tag{4.6}$$

Proof. It can be easily shown that under Condition (4.3), π_i^R, $i = 1, 2$, do not depend on the choices of $\pi_2^{(0)}(\boldsymbol{\phi})$ in *Step 0*. Hence the marginal reference priors for θ and $\boldsymbol{\phi}$ are formed of (4.5), and the independent reference prior for $(\theta, \boldsymbol{\phi})$ is given by (4.4). □

Ghosal (1999) investigated first order matching priors in non-regular cases. He considered the families of discontinuous densities used by Ghosal and Samanta (1997) and Ghosal (1997). He derived two differential equations which a first order matching prior for θ and a first order matching prior for $\boldsymbol{\phi}$ should satisfy, respectively. Now a sufficient condition under which the independent reference prior, given by (4.4), is a first order matching prior is provided.

Theorem 4.2 *Suppose*

$$\begin{aligned} c(\theta, \boldsymbol{\phi}) &= c_1(\theta)c_2(\boldsymbol{\phi}), \\ |\boldsymbol{\lambda}(\theta, \boldsymbol{\phi})| &= \lambda_1(\theta)\lambda_2(\boldsymbol{\phi}), \\ |\boldsymbol{\gamma}(\theta, \boldsymbol{\phi})| &\propto c(\theta, \boldsymbol{\phi})|\boldsymbol{\lambda}(\theta, \boldsymbol{\phi})|^{1/2} = c_1(\theta)c_2(\boldsymbol{\phi})\sqrt{\lambda_1(\theta)\lambda_2(\boldsymbol{\phi})}, \end{aligned} \tag{4.7}$$

where

$$\begin{aligned} c(\theta, \boldsymbol{\phi}) &= E\left\{\frac{\partial}{\partial \theta} \log f(X; \theta, \boldsymbol{\phi})\right\}, \\ \boldsymbol{\lambda}(\theta, \boldsymbol{\phi}) &= -E\left\{\frac{\partial^2}{\partial \varphi_j \partial \varphi_k} \log f(X; \theta, \boldsymbol{\phi})\right\}, \quad j, k = 1, \ldots, d, \\ \boldsymbol{\gamma}(\theta, \boldsymbol{\phi}) &= E\left\{\frac{\partial^2}{\partial \theta \partial \varphi_j} \log f(X; \theta, \boldsymbol{\phi})\right\}, \quad j = 1, \ldots, d. \end{aligned}$$

Then the independent reference prior $\pi^R(\theta, \boldsymbol{\phi})$ for $(\theta, \boldsymbol{\phi})$, given by (4.4), is a first order matching prior for θ and $\boldsymbol{\phi}$.

Proof. By Ghosal (1999), a first order matching prior $\pi(\theta, \boldsymbol{\phi})$ satisfies the following differential equation when θ is the parameter of interest,

$$\frac{1}{c(\theta, \boldsymbol{\phi})}\frac{\partial}{\partial \theta} \log \pi(\theta, \boldsymbol{\phi}) + \frac{|\boldsymbol{\gamma}(\theta, \boldsymbol{\phi})|}{c(\theta, \boldsymbol{\phi})|\boldsymbol{\lambda}(\theta, \boldsymbol{\phi})|}\frac{\partial}{\partial \boldsymbol{\phi}} \log \pi(\theta, \boldsymbol{\phi})$$

$$= -\frac{\partial}{\partial\theta}\left\{\frac{1}{c(\theta,\boldsymbol{\phi})}\right\} - \frac{\partial}{\partial\boldsymbol{\phi}}\left\{\frac{|\boldsymbol{\gamma}(\theta,\boldsymbol{\phi})|}{c(\theta,\boldsymbol{\phi})|\boldsymbol{\lambda}(\theta,\boldsymbol{\phi})|}\right\}. \tag{4.8}$$

When $\boldsymbol{\phi}$ is the interest parameter, a matching prior $\pi(\theta,\boldsymbol{\phi})$ satisfies

$$\frac{1}{|\boldsymbol{\lambda}(\theta,\boldsymbol{\phi})|^{1/2}}\frac{\partial}{\partial\boldsymbol{\phi}}\log\pi(\theta,\boldsymbol{\phi}) = -\frac{\partial}{\partial\boldsymbol{\phi}}\left\{\frac{1}{|\boldsymbol{\lambda}(\theta,\boldsymbol{\phi})|^{1/2}}\right\}. \tag{4.9}$$

We thus need to prove that the independent reference prior $\pi^R(\theta,\boldsymbol{\phi})$ for $(\theta,\boldsymbol{\phi})$, given by (4.4), is a solution to the equations, given by (4.8) and (4.9), under Condition (4.7). It is shown that

$$\begin{aligned}
&\frac{1}{c(\theta,\boldsymbol{\phi})}\frac{\partial}{\partial\theta}\log\pi^R(\theta,\boldsymbol{\phi}) + \frac{|\boldsymbol{\gamma}(\theta,\boldsymbol{\phi})|}{c(\theta,\boldsymbol{\phi})|\boldsymbol{\lambda}(\theta,\boldsymbol{\phi})|}\frac{\partial}{\partial\boldsymbol{\phi}}\log\pi^R(\theta,\boldsymbol{\phi}) \\
\propto\;& \frac{1}{c_1(\theta)c_2(\boldsymbol{\phi})}\frac{\partial}{\partial\theta}\log\left[c_1(\theta)\sqrt{\lambda_2(\boldsymbol{\phi})}\right] + \frac{c_1(\theta)c_2(\boldsymbol{\phi})\sqrt{\lambda_1(\theta)\lambda_2(\boldsymbol{\phi})}}{c_1(\theta)c_2(\boldsymbol{\phi})\lambda_1(\theta)\lambda_2(\boldsymbol{\phi})}\frac{\partial}{\partial\boldsymbol{\phi}}\log\left[c_1(\theta)\sqrt{\lambda_2(\boldsymbol{\phi})}\right] \\
=\;& \frac{1}{c_1(\theta)^2c_2(\boldsymbol{\phi})} + \frac{1}{2\sqrt{\lambda_1(\theta)}\lambda_2(\boldsymbol{\phi})^{3/2}}
\end{aligned}$$

and

$$\begin{aligned}
&-\frac{\partial}{\partial\theta}\left\{\frac{1}{c(\theta,\boldsymbol{\phi})}\right\} - \frac{\partial}{\partial\boldsymbol{\phi}}\left\{\frac{|\boldsymbol{\gamma}(\theta,\boldsymbol{\phi})|}{c(\theta,\boldsymbol{\phi})|\boldsymbol{\lambda}(\theta,\boldsymbol{\phi})|}\right\} \\
\propto\;& -\frac{\partial}{\partial\theta}\left\{\frac{1}{c_1(\theta)c_2(\boldsymbol{\phi})}\right\} - \frac{\partial}{\partial\boldsymbol{\phi}}\left\{\frac{c_1(\theta)c_2(\boldsymbol{\phi})\sqrt{\lambda_1(\theta)\lambda_2(\boldsymbol{\phi})}}{c_1(\theta)c_2(\boldsymbol{\phi})\lambda_1(\theta)\lambda_2(\boldsymbol{\phi})}\right\} \\
=\;& \frac{1}{c_1(\theta)^2c_2(\boldsymbol{\phi})} + \frac{1}{2\sqrt{\lambda_1(\theta)}\lambda_2(\boldsymbol{\phi})^{3/2}}.
\end{aligned}$$

Thus $\pi^R(\theta,\boldsymbol{\phi})$ satisfies the equation (4.8). Now it is also seen that

$$\begin{aligned}
\frac{1}{|\boldsymbol{\lambda}(\theta,\boldsymbol{\phi})|^{1/2}}\frac{\partial}{\partial\boldsymbol{\phi}}\log\pi^R(\theta,\boldsymbol{\phi}) &\propto \frac{1}{\sqrt{\lambda_1(\theta)\lambda_2(\boldsymbol{\phi})}}\frac{\partial}{\partial\boldsymbol{\phi}}\log\left[c_1(\theta)\sqrt{\lambda_2(\boldsymbol{\phi})}\right] \\
&= \frac{1}{2\sqrt{\lambda_1(\theta)}\lambda_2(\boldsymbol{\phi})^{3/2}}
\end{aligned}$$

and

$$\begin{aligned}
-\frac{\partial}{\partial\boldsymbol{\phi}}\left\{\frac{1}{|\boldsymbol{\lambda}(\theta,\boldsymbol{\phi})|^{1/2}}\right\} &= -\frac{\partial}{\partial\boldsymbol{\phi}}\left\{\frac{1}{\sqrt{\lambda_1(\theta)\lambda_2(\boldsymbol{\phi})}}\right\} \\
&= \frac{1}{2\sqrt{\lambda_1(\theta)}\lambda_2(\boldsymbol{\phi})^{3/2}}.
\end{aligned}$$

Thus $\pi^R(\theta, \boldsymbol{\phi})$ is a solution to the equation (4.9). Hence the independent reference prior $\pi^R(\theta, \boldsymbol{\phi})$ for $(\theta, \boldsymbol{\phi})$, given by (4.4), satisfies both of the differential equations (4.8) and (4.9). The result then holds. $\square$

Remark 4.1 *The independent reference prior $\pi^R(\theta, \boldsymbol{\phi})$ for $(\theta, \boldsymbol{\phi})$, given by (4.4), is always a first order matching prior for $\boldsymbol{\phi}$ under Condition (4.3).*

4.2 Examples

Some examples of non-regular cases, which were considered by Ghosal (1997) and Ghosal (1999), are studied in this section.

4.2.1 Location-scale Family

The density of a location-scale family with unknown location parameter θ ($\in \mathbb{R}$) and scale parameter φ (> 0) is

$$f(x; \theta, \varphi) = \frac{1}{\varphi}\, f_0\left(\frac{x-\theta}{\varphi}\right), \quad x > \theta,$$

where $f_0(\cdot)$ is a strictly positive density on $[0, \infty)$. Then

$$\begin{aligned} c(\theta, \varphi) &= \frac{1}{\varphi}\, f_0(0+), \\ |\boldsymbol{\lambda}(\theta, \varphi)| &= \frac{1}{\varphi^2} \int \frac{(1 + x f_0'(x))^2}{f_0(x)} dx, \\ |\boldsymbol{\gamma}(\theta, \varphi)| &= \frac{2a}{\varphi^2} \ \text{ for some constant } a. \end{aligned}$$

Thus the marginal reference priors for θ and φ are

$$\pi_1^R(\theta) \propto 1, \ \theta \in \mathbb{R}, \tag{4.10}$$

$$\pi_2^R(\varphi) \quad \propto \quad \frac{1}{\varphi},\ \varphi > 0, \tag{4.11}$$

and the independent reference prior for (θ, φ) is

$$\pi^R(\theta, \varphi) \propto \frac{1}{\varphi}. \tag{4.12}$$

It is a first order matching prior for θ and φ, which is the same result as Ghosal (1999). By Ghosal (1997), the independent reference prior for (θ, φ) is the reference prior for (θ, φ) when one of the parameters θ or φ is the parameter of interest and the other is nuisance parameter.

4.2.2 Truncated Weibull Model

Consider the Weibull distribution with known shape parameter α (> 0), unknown scale parameter φ (> 0), and truncated at the left at some unknown point θ (> 0). Then for $x > \theta$, the density is

$$f(x; \theta, \varphi) = \alpha\varphi^\alpha x^{\alpha-1} \exp[-\varphi^\alpha(x^\alpha - \theta^\alpha)].$$

Thus

$$\begin{aligned} c(\theta, \varphi) &= \alpha\theta^{\alpha-1}\varphi^\alpha, \\ |\boldsymbol{\lambda}(\theta, \varphi)| &= \frac{\alpha^2}{\varphi^2}, \\ |\boldsymbol{\gamma}(\theta, \varphi)| &= 2\alpha^2\theta^{\alpha-1}\varphi^{\alpha-1}. \end{aligned}$$

Hence the marginal reference priors for θ and φ are

$$\pi_1^R(\theta) \quad \propto \quad \theta^{\alpha-1},\ \theta > 0, \tag{4.13}$$

$$\pi_2^R(\varphi) \quad \propto \quad \frac{1}{\varphi},\ \varphi > 0, \tag{4.14}$$

and the independent reference prior for (θ, φ) is

$$\pi^R(\theta, \varphi) \propto \frac{\theta^{\alpha-1}}{\varphi}. \tag{4.15}$$

It is also a first order matching prior for θ and φ. Ghosal (1999) also obtained the same first order matching priors. The independent reference prior for (θ, φ) is equivalent to the reference prior for (θ, φ), derived by Ghosal (1997), when one of the parameters θ or φ is the interest parameter and the other is nuisance parameter.

Chapter 5

Summary and Future Work

In Bayesian inference, sufficient information on priors is not often available. Then objective priors could be a good choice instead of subjective priors. Thus developing objective priors has been of great interest in Bayesian methdology. There are various objective priors other than a constant prior; for example, the Jeffreys prior, a reference prior, an independent reference prior and a probability matching prior.

We studied a class of objective (noninformative) priors based on the independent reference prior which was introduced by Sun and Berger (1998). An independent reference prior is derived by assuming that the groups of parameters are independent. Most of the reference priors have the independence property in the sense that they formed as the product of marginal reference priors. Independent reference priors do not depend on the order of inferential importance of the parameters while reference priors definitely do. In practice, partial information on priors such as the independence of the parameters might be available. Hence, in real applications of Bayesian inference, an independent reference prior could be used.

In Chapter 2, we extended and generalized the independent reference prior by considering multiple groups of parameters while Sun and Berger (1998) used two groups of parameters. An iterative algorithm to derive the general independent reference prior was given first. A mild sufficient condition under which we obtain a closed form of the independent reference prior without going through the iterations was then provided. Two useful results from the independent reference prior were provided. First, the invariance of the independent reference prior was proven under a type of one-to-one reparameterization where the Jacobian matrix is diagonal. Second, it was shown that the independent reference prior is a first order matching prior under a sufficient condition. In Chapter 3, the independent reference priors were derived for various examples. It turned out that they are identical to the reference priors in most cases. It was also observed that the independent reference priors are the first order matching priors in most of the examples. In Chapter 4, we provided an iterative algorithm to obtain an independent reference prior for some types of non-regular cases where the support of the data is monotonically decreasing or increasing in a non-regular type parameter. A sufficient condition under which a closed form of the independent reference prior is derived was established. It was proven that the independent reference prior is a first order matching prior under a sufficient condition. Some examples were also given.

In most of the examples in Chapter 3, the sufficient condition, given by (2.3), in Theorem 2.1 and the sufficient condition, given by (2.8), in Theorem 2.3 were satisfied. Thus we obtained the closed forms of independent reference priors which were also the first order matching priors. However in the five examples the conditions were not satisfied. The independent reference priors were derived by using the iterative algorithm for such examples. They are given in Proposition 3.2, 3.4 and 3.5–3.7. The marginal reference priors were proper in Proposition 3.2 and 3.4 so that we did not use compact sets in the algorithm. However we implemented the algorithm using compact sets in Proposition 3.5–3.7 since there were improper marginal reference priors. In Proposition 3.5 and 3.6, we obtained the closed form of the independent reference prior by choosing a specific sequence of compact sets for an improper marginal reference prior. In Proposition 3.7, it was practically hard to derive a closed form of the independent reference prior since the integration on a sequence of compact sets was impossible. Thus choice of a sequence of compact sets for an improper marginal reference prior or calculation with respect to a sequence of compact sets could be an issue. Hence we might need to define an explicit expression for the independent reference prior which excludes the iterations and any compact set operations as Berger, Bernardo and Sun (2007) did for the reference prior.

We derived and investigated an independent reference prior for regular cases where the data has common support and the Fisher information matrix is available. All of the examples considered in Chapter 3 are apparently the regular cases. In Chapter 4, we derived and studied an independent reference prior only for some types of non-regular cases where the support of the data is either monotonically decreasing or increasing in a non-regular type parameter. Thus our current results do not include all the cases. Hence the results for more general cases need to be developed.

Bibliography

Bayarri, M. J., Berger, J. O. and Datta, G. S. (2007). Objective Bayes testing of Poisson versus inflated Poisson models. preprint.

Berger, J. O. (1992). Discussion of "Non-informative priors (with discussion)" by J. K. Ghosh and R. Mukerjee. *Bayesian Statistics, 4*, Eds. J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, pp. 205-206, Oxford University Press, Oxford.

Berger, J. O. and Bernardo, J. M. (1989). Estimating a product of means: Bayesian analysis with reference priors. *Journal of the American Statistical Association*, **84**, 200-207.

Berger, J. O. and Bernardo, J. M. (1992). On the development of reference priors (with discussion). *Bayesian Statistics, 4*, Eds. J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, pp. 35-60, Oxford University Press, Oxford.

Berger, J. O., Bernardo, J. M. and Sun, D. (2007). The formal definition of reference priors. *The Annals of Statistics*, in press.

Berger, J. O. and Sun, D. (2007). Objective priors for the bivariate normal model. *The Annals of Statistics*, *The Annals of Statistics*, in press.

Bernardo, J. M. (1977). Inference about the ratio of normal means: a Bayesian approach to the Fieller-Creasy problem. *Recent Developments in Statistics*, Eds. J. R. Barra et al., pp. 345-350, North-Holland, Amsterdam.

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference (with discussion). *Journal of the Royal Statistical Society, Series B*, **41**, 113-147.

Crowder, M. and Sweeting, T. (1989). Bayesian inference for a bivariate binomial distribution. *Biometrika*, **82**, 37-45.

Datta, G. S. and Ghosh, J. K. (1995). On priors providing frequentist validity for Bayesian inference. *Biometrika*, **76**, 599-603.

Datta, G. S. and Ghosh, M. (1995). Some remarks on noninformative priors. *Journal of the American Statistical Association*, **90**, 1357-1363.

Datta, G. S. and Ghosh, M. (1996). On the invariance of noninformative priors. *The Annals of Statistics*, **24**, 141-159.

Datta, G. S. and Mukerjee, R. (2004). *Probability Matching Priors: Higher Order Asymptotics.* Springer-Verlag, New York.

Ferreira, M. A. R. (2007). Objective Bayesian analysis for the Student-t regression model. *The Sixth International Workshop on Objective Bayes Methodology*, Rome, Italy.

Garvan, C. W. and Ghosh, M. (1997). Noninformative priors for dispersion models. *Biometrika*, **84**, 976-982.

Ghosal, S. (1997). Reference priors in multiparameter nonregular cases. *Test*, **6**, 159-186.

Ghosal, S. (1999). Probability matching priors for non-regular cases. *Biometrika*, **86**, 956-964.

Ghosal, S. and Samanta, T. (1997). Expansion of Bayes risk for entropy loss and reference prior in nonregular cases. *Statistics & Decisions*, **15**, 129-140.

Ghosh, J. K. and Mukerjee, R. (1992). Non-informative priors (with discussion). *Bayesian Statistics, 4*, Eds. J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, pp. 195-210, Oxford University Press, Oxford.

Jeffreys, H. (1961). *Theory of Probability.* Oxford University Press, Oxford.

Laplace, P. (1812). *Théorie Analytique des Probabilités.* Courcier.

Sun, D. (1997). A note on noninformative priors for Weibull distributions. *Journal of Statistical Planning and Inference*, **61**, 319-338.

Sun, D. and Berger, J. O. (1998). Reference priors with partial information. *Biometrika*, **85**, 55-71.

Sun, D., Ghosh, M. and Basu, A. P. (1998). Bayesian analysis for a stress-strength system under noninformative priors. *The Canadian Journal of Statistics*, **26**, 323-332

Sun, D. and Ye, K. (1995). Reference prior Bayesian analysis for normal mean products. *Journal of the American Statistical Association*, **90**, 589-597.

Sun, D. and Ye, K. (1996). Frequentist validity of posterior quantiles for a two-parameter exponential family. *Biometrika*, **83**, 55-65.

Sun, D. and Ye, K. (1999). Reference priors for a product of normal means when variances are unknown. *The Canadian Journal of Statistics*, **27**, 97-103.

Welch, B. L. and Peers, H. W. (1963). On formulae for confidence points based on intervals of weighted likelihoods. *Journal of the Royal Statistical Society, Series B*, **25**, 318-329.

Yang, R. and Berger, J. O. (1997). A catalog of noninformative priors. *Discussion Paper, 1997-42*, ISDS, Duke University, Durham, NC.

Vita

Mi Hyun Lee was born in Seoul, Republic of Korea. She received her Bachelor of Science degree and Master of Science degree in Statistics from the department of Statistics, Ewha Womans University, Seoul, Korea in February, 1999 and 2001, respectively. She then earned her Master of Statistics degree from the department of Statistics, North Carolina State University in May, 2004. She completed her Doctor of Philosophy degree in Statistics in the department of Statistics, Virginia Polytechnic Institute and State University in December, 2007. The author is a member of the National Statistics Honorary Society (Mu Sigma Rho).