Chapter 5

Using Copulas for Bayesian Meta-Analysis

5.1 Introduction

For Bayesian meta-analysis of treatment effectiveness in medical sciences, a given treatment is applied to \( N \) samples of patients where \( n_i \) is the size of \( i^{th} \) sample for \( i = 1, 2, ..., N \). For \( i^{th} \) sample, the sample effectiveness is \( x_i \) where \( x_i = (x_{i1}, ..., x_{in_i}) \) is assumed to be normally distributed \( N(n_i|x_i, \theta_i, \sigma^2 I_{n_i}) \). Here conditional on \( i^{th} \) sample, \( \theta_i \) represents the treatment effectiveness and \( \sigma^2 \) is the variance. Non-observable meta-variable \( x \) follows \( N(x | \theta, \sigma^2) \) where the meta-parameter \( \theta \) is the unconditional treatment effectiveness and \( \sigma^2 \) is the meta-variance.

In most of the situations, the sample variables \( x_i \) may be discrete denoting the number of successes. Hence the normal hierarchical model is applied to the logit transformation \( y_i = \log(x_i/(n_i - x_i)) \) which can be reparametrized as \( \log(\theta_i/(1 - \theta_i)) \). When the multicentre binomial sample contains zeros, this normal approximation does not work properly, even by applying a continuity correction to the original data (J Sweeting et al. (2004)).

This chapter proposes a Bayesian model for meta-analysis of multicentre Binomial data containing zeros in which no continuity correction is needed. The proper objective priors for the Binomial parameter \( \theta_i \) and the meta-parameter \( \theta \) are used (Lambert et al. (2005)).

The selection of the prior distribution used for linking \( \theta_i \) and \( \theta \) is a crucial step in Bayesian meta-analysis. The choice of the distribution of \( \theta_i \) conditional on \( \theta \), is made by putting two restrictions on it. These conditions are that the linking distribution

(i) gives specified marginals,

(ii) accommodates different degrees of heterogeneity between the experimental centres.

There is vast literature available for constructing bi-dimensional distributions \( \pi(\theta_i, \theta) \) with fixed marginals \( \pi(\theta_i) \) and \( \pi(\theta) \), which is known as the Frechet problem. The class of bivariate distributions with fixed marginals is the Frechet class. Relevant subclasses of the Frechet class
are those given by Frechet-Hoeffding, Farlie-Gumbel-Morgenstern (FGM), Clayton-Oakes, Sarmanov, Ali-Mikhail-Haq and Cuadras-Auge, among many others (Joe (1997)). All of these models assume a particular dependence structure between \( \theta_i \) and \( \theta \). The interest is in choosing a subclass of the Frechet class accommodating a wide range of heterogeneity degrees between centres. The parametric subclass of the Frechet class proposes satisfies

(i) the between center heterogeneity, that is, the variance of \( \theta_i \) can be parameterized in terms of the Pearson’s correlation coefficient between \( \theta_i \) and \( \theta \),

(ii) as the prior in the subclass varies, the correlation coefficient lies in \((0,1)\).

The former constraint considerably simplifies the task of choosing priors for the hyperparameters for integrating them out. The latter constraint entails considering the union of several parametric subclasses of the Frechet class which enter in the model as a new parameter. The Bayesian inference is made as a mixture of the conditional inference on the subclasses using Bayesian model averaging. Few references for meta-analysis using priors in the Frechet class are by Casella and Moreno (2005), Chu et al. (2010) and Chen et al. (2015).

A test for testing the equality of the meta-parameters of two alternative treatments is also proposed. The predictive distribution of the meta-variable \( z \) gives the distribution of the effect of the treatment for a new patient (S (1992)) which requires the posterior distribution of the meta-parameter \( \theta \).

The chapter is organized as follows. Section 5.2 gives brief introduction to theory of Copulas. In Section 5.3, we present the linking distribution of \( \theta_i \) and \( \theta \) for the Uniform and the Jeffrey priors and study some of their properties. The likelihood and the posterior distribution of the meta-parameter have been given. In Section 5.4, a Bayesian procedure for testing the equality of the meta-effectiveness of two treatments is given. In Section 5.5, the Bayesian procedure for meta-analysis is applied to four real data sets taken from Fisher and Van Belle (1993), Bellamy et al. (2009), (Collins et al. (1995); Eggar and Smith (1995); Egger et al. (1997)) and Niël-Weise et al. (2007). Section 5.6 contains concluding remarks.
5.2 Introduction to Copulas

According to NL (1997), Copulas are of interest to statisticians as a way of studying scale-free measures of dependence and as a starting point for constructing families of bivariate distributions. They are functions that join or couple multivariate distribution functions to their one dimensional marginal distribution functions. They are also looked upon as distribution functions with one-dimensional marginals as uniform (Nelsen (2007)). By using copulas, a model for the meta-analysis of diagnostic accuracy studies simply fits a bivariate distribution. This is done by differentiating the copula, considering the resulting density as the likelihood function and using standard maximum likelihood estimation methods for parameter estimation.

5.2.1 Definition of Copulas

Consider a pair of random variables $X$ and $Y$ with distribution functions $F(x)$ and $G(y)$, respectively and a joint distribution function $H(x, y)$. To each pair of real numbers $(x, y)$, three functions: $F(x)$, $G(y)$ and $H(x, y)$ can be associated where each function lying in the interval $[0, 1]$. In other words, each pair $(x, y)$ of real numbers leads to a point $(F(x), G(y))$ in the unit square $[0, 1] \times [0, 1]$ and this in turn corresponds to a function $H(x, y)$ with support on $[0, 1]$.

Copula is a function

- assigning the value of the joint distribution function to each ordered pair of values of the individual distribution functions,
- describing the dependence structure,
- containing all the information to link the marginal distribution to the joint distribution,
- helps in combining many marginals distribution functions to get a valid multivariate distribution function.

Nelsen et al. (1997) stated that copulas with quadratic section (Quesada-Molina and Rodríguez-Lallena (1995)) are not able to model large dependences. Hence, Copulas with cubic sections are introduced leading to the conclusion that copulas with higher order polynomial sections increase the degree of dependence and also the complexity of the model.
Definition:
A two dimension copula is a function \( C(x, y) \) from \([0, 1] \times [0, 1]\) to \([0, 1]\) with the properties:

- \( C(x, 0) = 0 = C(0, y) \),
- \( C(x, 1) = x \) and \( C(1, y) = y \),
- For every \( x_1, x_2 \) and \( y_1, y_2 \) such that \( 0 \leq x_1 < x_2 \leq 1 \) and \( 0 \leq y_1 < y_2 \leq 1 \), we have
  \[
  C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0.
  \]

A bivariate copula defined on the unit square \( I = [0, 1]^2 \) is a bivariate cumulative distribution function (cdf) with univariate marginals.

The result justifying the use of copulas for building bivariate distributions was given by Sklar (1959) in the form of Sklar’s Theorem which states that:

If \( F(x, y) \) is a joint distribution function with continuous marginals \( F_X(x) \) and \( F_Y(y) \), then there exists a unique copula \( C \) such that

\[
F(x, y) = C(F_X(x), F_Y(y)).
\]

To quantify the correlation between \( X \) and \( Y \), it was shown by Quesada-Molina and Rodriguez-Lallena (1995) that both Spearman’s \( \rho_S(X, Y) \) and Kendall’s \( \tau(X, Y) \) correlation coefficients can be described in terms of the copula as

\[
\rho_S(X, Y) = 12E\{(F_X(x) - 1/2)(F_Y(y) - 1/2)\}
\[
= 12 \int \int \{C(u,v) - uv\} \, dudv
\]

and

\[
\tau(X, Y) = P\{(X - X^*)(Y - Y^*) > 0\} - P\{(X - X^*)(Y - Y^*) < 0\}
\]

\[
= 4 \int \int C(u,v) \, dC(u,v) - 1
\]

where \( (X^*, Y^*) \) is an independent copy of \( (X, Y) \). It should be noted that in general, the Pearson’s correlation between \( X \) and \( Y \) depends also on the marginal distributions and is thus affected by changes in scale.

In literature, there are many parametric copula families available whose parameters control
the strength of dependence. One important parametric family of copulas is the FGM (Farlie-Gumbel-Morgenstern) family (Moreno et al. (2014)).

5.3 Bayesian Models for Sparse Binomial Data

When a clinical trial is carried out in \( N \) different healthcare centres, the quantity of interest is often the effectiveness of the new treatment which is non-observable. This has to be derived from the performance of the new treatment in the healthcare centres for which data sets are available.

For \( i = 1, 2, \ldots, N \), let

\[
x_i = \begin{cases} 
1 & \text{if the person receiving the treatment at centre } i \text{ succeeds} \\
0 & \text{otherwise}
\end{cases}
\]

and \( \theta_i \) be the probability of success of treatment.

Let

\[
M_i : \{Ber(x_i|\theta_i), \pi(\theta_i)\}, \quad i = 1, 2, \ldots, N, \tag{5.3.1}
\]

be the Bayesian model for the observational study in centre \( i \) where \( Ber(x_i|\theta_i) = \theta_i^{x_i}(1-\theta_i)^{1-x_i} \).

The prior will be either the uniform prior \( \pi^U(\theta_i) = Beta(\theta_i|1, 1) \) or the Jeffrey prior given by \( \pi^J(\theta_i) = Beta(\theta_i|1/2, 1/2) \). Uniform prior was considered by Bayes and Laplace and Jeffery prior is invariance under one to one reparametrization. These priors are widely accepted as objective priors for the Bernoulli parameter. For sparse data, Tuyl et al. (2008) give some arguments favouring Uniform over Jeffrey prior. Similarly, the Bayesian meta-model for unobservable meta-treatment effectiveness \( x \) is given by the pair of distributions

\[
M_0 : \{Ber(x|\theta), \pi(\theta)\} \tag{5.3.2}
\]

where \( x \) takes values 0 or 1;

\( \theta \), the meta parameter, is the unconditional probability of success of the treatment;

\( \pi(\theta) \) can be Uniform or Jeffrey’s prior.
5.3.1 Linking Distribution

A linking distribution $\pi(\theta | \theta_i)$ with fixed marginals $\pi(\theta_i)$ and $\pi(\theta)$ is needed for formulation of Bayesian model. For this $\pi(\theta_i | \theta) = \pi(\theta | \theta_i) \pi(\theta)$ must satisfy the integral equations

$$ \int_0^1 \pi(\theta_i, \theta) d\theta_i = \pi(\theta), \quad \int_0^1 \pi(\theta, \theta_i) d\theta = \pi(\theta_i) \quad (5.3.3) $$

5.3.2 The Uniform Marginals

For uniform marginals, $\pi(\theta) = 1_{(0,1)}(\theta)$ and $\pi(\theta_i) = 1_{(0,1)}(\theta_i)$, a class of bivariate distributions is the FGM copula (Moreno et al. (2014)) given by

$$ C_k^{FGM}(\theta_i, \theta) = \theta_i \theta + k \theta_i (1 - \theta)(1 - \theta) \quad \theta_i \geq 0, \theta \leq 1 \quad \text{and} \quad 0 \leq k \leq 1 \quad (5.3.4) $$

The corresponding FGM parametric class of probability distribution satisfying (5.3.3) is written as

$$ \pi^{FGM}(\theta_i, \theta | k) = 1 + k(2\theta_i - 1)(2\theta - 1). \quad (5.3.5) $$

5.3.3 Limitations of FGM Family

- The between-center heterogeneity captured by this class is not sufficiently wide. This heterogeneity is measured by the variance of $\theta_i$ conditional on $\theta$ and $k$ and is given by

$$ V^{FGM}(\theta_i | \theta, k) = \frac{1}{3} + \frac{k}{6} (2\theta - 1) - \left( \frac{1}{2} + \frac{k}{6} (2\theta - 1) \right)^2 $$

which is a decreasing function of $k$ for any $\theta$.

- Since $\rho = \frac{k}{3}$, it follows that $V^{FGM}(\theta_i | \theta, k)$ is a decreasing function of the correlation coefficient $\rho$.

- $0 \leq \rho \leq \frac{1}{3}$ as $k$ varies in $(0, 1)$. Hence the FGM class is unable to model a class with correlation coefficient between $\theta_i$ and $\theta$, greater than $\frac{1}{3}$.

- Although the FGM model is an interesting family constructed from specified marginals, this model cannot be used to represent the joint distribution of two highly correlated variables.
5.3 Bayesian Models for Sparse Binomial Data

Assuming a uniform prior for \( \rho \) in \((0, 1/3)\), the induced distribution on \( k \) is uniform in \((0, 1)\) and the linking distribution of FGM is given by

\[
\pi^{FGM}(\theta_i|\theta) = 1 + \frac{1}{2}(2\theta_i - 1)(2\theta - 1).
\]  

(5.3.6)

To increase dependence between variables, Bairamov and Kotz (2002) extended the bivariate FGM distribution.

The main features of this family are that

- it increases the dependence between the underlying variables by introducing additional parameters;

- It allows us to achieve correlation more than 0.5.

Generalized FGM copula (Bairamov and Kotz (2002)) is defined as

\[
C_k^{FGM}(\theta_i, \theta) = \theta_i\theta[1 + k(1 - \theta_i)^p(1 - \theta)^p] \quad p > 0, \quad q > 0, \quad \theta_i > 0, \quad \theta < 1
\]  

(5.3.7)

The corresponding parametric class of probability distribution is given by

\[
\pi^{GFGM}(\theta_i, \theta|k, p) = 1 + k(1 - \theta_i)^{p-1}(1 - \theta)^{p-1}[1 - \theta_i^q(1 + qp)][1 - \theta^q(1 + qp)].
\]  

(5.3.8)

The range for \( k \) is given by

\[
-\min \left\{ \frac{1}{q^2} \left( \frac{1 + qp}{q(p - 1)} \right)^{2(p-1)}, 1 \right\} \leq k \leq \frac{1 + qp}{q(p - 1)}^{p-1}.
\]

The range of \( \rho \) for a specific value of \( p \) is given by

\[
\frac{-3}{(p + 1)^2} \min \left\{ \frac{1}{4} \left( \frac{1}{1 + \frac{3q}{p+3p}} \right)^{2(p-1)}, 1 \right\} \leq \rho \leq \frac{3}{2(p + 1)^2} \left( \frac{1}{1 + \frac{3q}{p+3p}} \right)^{p-1}.
\]  

(5.3.9)

It is observed from (5.3.9) that \( \rho \to 0 \) as \( p \to \infty \).

It implies that for large values of \( p \), there is weak dependence between the components \( \theta_i \) and \( \theta \).

A refined iterative search yields

\[
\rho_{max}(p) = 0.501595 \text{ at } q = 2.89, \ p = 1.50 \text{ and }
\]

\[
\rho_{min}(p) = -0.48 \text{ at } q = 2, \ p = 1.50 \ (\text{Bairamov and Kotz (2002)}).
\]
5.3 Bayesian Models for Sparse Binomial Data

The fact that $\rho_{\max}(p)$ exceeds $1/2$ is advantageous for applications purpose.

We use the probability distribution given by (5.3.8) with $q = 3$

$$\pi^{GFGM}(\theta_i, \theta|k, p) = 1 + k(1 - \theta_i^2)^{p-1}(1 - \theta^2)^{p-1}[1 - \theta_i^2(1 + 3p)][1 - \theta^2(1 + 3p)]. \quad (5.3.10)$$

The Linking distribution which we use with $q = 3$ is given by

$$\pi^{GFGM}(\theta_i|\theta, k, p) = \frac{\pi^{GFGM}(\theta_i, \theta|k, p)}{\pi(\theta|k, p)}$$

$$= 1 + k(1 - \theta_i^2)^{p-1}(1 - \theta^2)^{p-1}[1 - \theta_i^2(1 + 3p)][1 - \theta^2(1 + 3p)]$$

(5.3.11)

where $\pi(\theta|k, p)$ is probability function of $B(1, 1)$.

It is observed that

$$E^{GFGM}(\theta_i|\theta, k, p) = \frac{1}{2} + \frac{k(1 - \theta^2)^{p-1}[1 + \theta^2(1 + 3p)] \text{Beta}(2, 1 + p)}{3}$$

$$E^{GFGM}(\theta_i^2|\theta, k, p) = \frac{1}{3} + \frac{2k(1 - \theta^2)^{p-1}[1 + \theta^2(1 + 3p)]}{3(1 + p)}$$

Hence

$$V^{GFGM}(\theta_i|\theta, k, p) = \frac{1}{3} + \frac{2k(1 - \theta^2)^{p-1}[1 + \theta^2(1 + 3p)]}{3(1 + p)} - \left[ \frac{1}{2} + \frac{k(1 - \theta^2)^{p-1}[1 + \theta^2(1 + 3p)] \text{Beta}(2, 1 + p)}{3} \right]^2.$$

The conditional variance $V^{GFGM}(\theta_i|\theta, k, p)$ is a decreasing function of $\rho$.

5.3.4 The Jeffrey Marginals

If the Jeffrey priors $\pi^J(\theta_i)$ and $\pi^J(\theta)$ are used instead of the Uniform priors, then the Sarmanov-Jeffreys (SJ) family (Moreno et al. (2014)) is considered. The probability distribution of SJ family is written as

$$\pi^{SJ}(\theta_i, \theta|k) = \frac{1}{\pi^2}(\theta_i(1 - \theta_i)(1 - \theta))^{-1/2} \left( 1 + \frac{k}{4}(2\theta_i - 1)(2\theta - 1) \right) \quad (5.3.12)$$
where $0 \leq k \leq 4$. This family satisfies (5.3.3) and the conditional variance is seen to be

$$V^{SJ}(\theta_i|\theta, k) = \frac{32 - k^2(2\theta - 1)^2}{256} \quad (5.3.13)$$

which is a decreasing function of correlation coefficient $\rho = \frac{k}{8}$.

Furthermore, as $k$ ranges over $(0, 4)$, we have $0 \leq \rho \leq \frac{1}{2}$. Hence this class needs to be enlarged to get a coefficient more than $1/2$.

Assuming a uniform prior for $\rho$ in $(0, 1/2)$, the induced distribution on $k$ is uniform in $(0, 4)$ and the linking distribution of $SJ$ is given by

$$\pi^{SJ}(\theta_i|\theta) = \frac{1}{\pi_2}(\theta_i(1 - \theta_i)(1 - \theta))^{-1/2} \left(1 + \frac{1}{2}(2\theta_i - 1)(2\theta - 1)\right). \quad (5.3.14)$$

Lin and Huang (2011) considered a generalized version of Sarmanov family, introduced earlier by Bairamov et al. (2001). They showed that the generalized Sarmanov family always has a correlation approaching one regardless of the marginals, as long as the marginals are of the same type. We consider generalization of Sarmanov and Sarmanov-Lee models constructed by Bairamov et al. (2011). These distributions have a simple analytical form like the FGM and as in the normal case, the correlation coefficient $\rho$ totally governs the dependence between the variables.

The Generalized Sarmanov copula is defined as

$$C^{GSJ}(\theta_i|\theta, k, p) = \theta_i + \frac{k}{p+1} \left\{ \theta_i^{p+1} - \theta_i^{p+1}(\theta + \theta_i^{p+1}) - \theta_i^{p+1}(\theta_i^{p+1} - \theta_i^{p+1}) \right\}, \quad \theta_i \geq 0, \quad \theta \leq 1, \quad p > 0 \quad (5.3.15)$$

with corresponding parametric class of probability distribution as

$$\pi^{GSJ}(\theta_i|\theta, k, p) = 1 + k \left\{ \theta_i^{p} + \theta_i^{p} - (p + 1)\theta_i^{p} + \frac{1}{p+1} \right\}. \quad (5.3.16)$$

For $p \geq 1$, $k$ satisfying

$$-\frac{p+1}{p} \leq k \leq \frac{p+1}{2(p+1) - (p+1)^2 - 1}.$$
and for $0 < p < 1$, $k$ satisfying

$$\frac{-p + 1}{p} \leq k \leq p + 1.$$  

Pearson’s correlation coefficient $\rho$ between $\theta$, and $\theta$ is obtained as

$$\rho = \frac{-3kp^2}{(p + 1)(p + 2)^2}.$$  

It is observed that

$$\frac{-p^2}{(p + 2)^2} \leq \rho \leq \frac{3p}{(p + 2)^2} \quad \text{for } 0 < p < 1,$$

and

$$\frac{-3}{(p + 2)^2} \leq \rho \leq \frac{3p}{(p + 2)^2} \quad \text{for } p \geq 1.$$  

The expressions for conditional expectations and variance are given below as

$$E^{GSJ}(\theta_1|\theta, k, p) = \pi \sqrt{\theta(1-\theta)} \left[ \frac{1 + k\theta}{2} - \frac{k[(p + 1)\theta - 1]}{p + 2} - \frac{k}{2(p + 1)} \right],$$  

$$E^{GSJ}(\theta_2|\theta, k, p) = \pi \sqrt{\theta(1-\theta)} \left[ \frac{1 + k\theta^p}{3} - \frac{k[(p + 1)\theta^p - 1]}{p + 3} - \frac{k}{3(p + 1)} \right],$$  

$$V^{GSJ}(\theta_1|\theta, k, p) = \frac{3\pi^2 k^2 p^2 \theta(1-\theta)}{\rho(p + 2)^2} \left\{ k(p + 2)^2 \theta^p (p + 1) - 1 \right\}$$

$$+ (p + 1)(k(5p^2 + 12p + 8) - p)$$

$$+ \theta^p(p + 4k(p + 1)\theta^p - k(p + 2)\theta^p) + k\theta^{p-1}(p + 1)(p + 2)$$

$$+ k(p + 1)^2(\theta^p(p^2 + 3p) - (p + 2)) \right\}$$

$$+ \pi \sqrt{\theta(1-\theta)} \left[ \frac{1 + k\theta^p}{3} + k \left\{ \frac{1 - (p + 1)\theta}{p + 3} - \frac{1}{3(p + 1)} \right\} \right] - \frac{\pi^2 \theta(1-\theta)}{4}.$$  

The conditional variance is a decreasing function of $\rho$. 
5.3.5 The Likelihood of the Meta-Parameter

For a given linking distribution \( \pi(\theta_i|\theta) \), the likelihood of the meta-parameter \( \theta \) for the data \((x_i, n_i)\) drawn from centre \(i\) is given by

\[
Pr(x_i|n_i, \theta) = \int_0^1 \binom{n_i}{x_i} \theta^{x_i}(1 - \theta)^{n_i-x_i} \pi(\theta_i|\theta) d\theta_i .
\]  

(5.3.17)

Replacing \( \pi(\theta_i|\theta) \) in (5.3.17) by the linking distributions (5.3.6), (5.3.11), (5.3.14) and (5.3.16), we obtain the conditional likelihoods as

\[
Pr(x_i|n_i, \theta, FGM) = \frac{1}{n_i + 1} \left( 1 + \frac{1}{2} (2\theta - 1) \left( \frac{2x_i - n_i}{n_i + 2} \right) \right),
\]

\[
Pr(x_i|n_i, \theta, GFGM) = \frac{n_i}{81} \Gamma(n_i - x_i + 1) \left[ \frac{81 \Gamma(1 + x_i)}{\Gamma(2 + n_i)} - 2 \times \pi \times k \times (1 - \theta^3)^{p-1}(-1 + \theta^3(1 + 3p)) \left\{ \frac{3(2.5-n_i)}{\Gamma(1 + n_i)} \right. \\
\left. - \frac{\Gamma(1 + n_i)}{\Gamma(2 + n_i)} \Gamma(1 + 3p) \right\} \frac{\Gamma(4 + x_i) \Gamma(1 + n_i)}{\Gamma(2 + n_i)} \right],
\]

where \( _3F_2(a, b, 1) \) and \( _3F_2(a_1, b_1, 1) \) denote the generalized hypergeometric functions with

\[a = \left( 1 - p, \frac{1 + x_i}{3}, \frac{2 + x_i}{3}, 1 + \frac{x_i}{3} \right), \quad b = \left( \frac{2 + n_i}{3}, \frac{1 + n_i}{3}, \frac{4 + n_i}{3} \right),\]

\[a_1 = \left( 1 - p, \frac{4 + x_i}{3}, \frac{5 + x_i}{3}, 2 + \frac{x_i}{3} \right), \quad b_1 = \left( \frac{5 + n_i}{3}, 2 + \frac{n_i}{3}, \frac{7 + n_i}{3} \right).\]

For the Sarmanov and Generalized Sarmanov class,

\[
Pr(x_i|n_i, \theta, SJ) = \binom{n_i}{x_i} Be(n_i - x_i + 1/2, x_i + 1/2) \frac{1 + n_i(3 - 4\theta) + 4x_i(2\theta - 1)}{\pi(n_i + 1)},
\]

and

\[
Pr(x_i|n_i, \theta, GSJ) = \frac{1}{(n_i + 1)} (1 + k\theta^p - \frac{k}{p + 1}) + (k - k\theta^p - kp\theta^p) \frac{\Gamma(n_i + 1) \Gamma(x_i + p + 1)}{\Gamma(x_i + 1) \Gamma(n_i + p + 2)}.
\]

Assuming that the centre samples \( \{(x_i, n_i), i = 1, ..., N\} \) are independent, the conditional
likelihood of the meta-parameter $\theta$ for the whole data set is written as

$$Pr(x|n, \theta, w) = \prod_{i=1}^{N} Pr(x_i|n_i, \theta, w)$$

(5.3.18)

where $x = (x_1, \ldots, x_N)$,

$n = (n_1, \ldots, n_N)$,

$\omega \in \{FGM, GFGM, SJ, GSJ\}$.

Assuming that the links under the uniform marginals are apriori equally likely, that is,

$Pr(FGM) = Pr(GFGM) = 1/2$, the conditional likelihood of the meta-parameter $\theta$ under the uniform marginals is given by

$$Pr^U(x|n, \theta) = \frac{1}{2} Pr^U(x|n, \theta, FGM) + \frac{1}{2} Pr^U(x|n, \theta, GFGM).$$

(5.3.19)

Similarly assuming apriori that $Pr(SJ) = Pr(GSJ) = 1/2$, the conditional likelihood of the meta-parameter $\theta$ under the Jeffrey marginals is

$$Pr^J(x|n, \theta) = \frac{1}{2} Pr^J(x|n, \theta, SJ) + \frac{1}{2} Pr^J(x|n, \theta, GSJ).$$

(5.3.20)

### 5.3.6 The Posterior Distribution of the Meta-Parameter

From the conditional likelihood $Pr^U(x|n, \theta)$ in (5.3.19) and prior $\pi(\theta) = \mathcal{I}_{(0,1)}(\theta)$, we have that the posterior distribution of the meta-parameter $\theta$ under the uniform marginals is given by

$$\pi^U(\theta|x, n) = \frac{Pr^U(x|n, \theta)}{\int_0^1 Pr^U(x|n, \theta) d\theta}.$$ 

(5.3.21)

We note that this is a mixture of the posterior probability of $\theta$ under FGM and GFGM, where the weights are the posterior probabilities of the Bayesian models for FGM and GFGM written as

$$M^{FGM} : \left\{ Pr(x|n, \theta, FGM), \pi(\theta, FGM) = \frac{1}{2} \mathcal{I}_{(0,1)}(\theta) \right\}$$

and

$$M^{GFGM} : \left\{ Pr(x|n, \theta, GFGM), \pi(\theta, GFGM) = \frac{1}{2} \mathcal{I}_{(0,1)}(\theta) \right\}.$$
The corresponding posterior model probabilities are given by

\[ Pr(M_{FGM} | x, n) = \frac{m(x|n, FGM)}{m(x|n, FGM) + m(x|n, GFGM)} \]
and

\[ Pr(M_{GFGM} | x, n) = 1 - Pr(M_{FGM} | x, n). \]

where

\[ m(x|n, FGM) = \int_0^1 Pr(x|n, \theta, FGM) d\theta \]
and

\[ m(x|n, GFGM) = \int_0^1 Pr(x|n, \theta, GFGM) d\theta . \]

are the likelihoods of FGM and GFGM respectively.

These weights indicate the relevance of the links on the posterior inference on \( \theta \).

Similarly, under the Jeffrey marginals, the posterior distribution of \( \theta \) is given by

\[ \pi^J(\theta|x, n) = \frac{Pr^J(x|n, \theta) \text{Beta}(\theta|1/2, 1/2)}{\int_0^1 Pr^J(x|n, \theta) \text{Beta}(\theta|1/2, 1/2) d\theta} . \] (5.3.22)

This posterior distribution can be seen as a mixture of the posterior probability of \( \theta \) under SJ and GSJ with weights as the posterior probabilities of the Bayesian models for SJ and GSJ.

Consider the Bayesian models for SJ and GSJ as

\[ M^{SJ} : \left\{ Pr(x|n, \theta, SJ), \pi(\theta, SJ) = \frac{1}{2} I_{(0,1)}(\theta) \right\} \]
and

\[ M^{GSJ} : \left\{ Pr(x|n, \theta, GSJ), \pi(\theta, GSJ) = \frac{1}{2} I_{(0,1)}(\theta) \right\} . \]

It is observed that weights in (5.3.22) are the posterior probabilities of \( M^{SJ} \) and \( M^{GSJ} \) and are given by

\[ Pr(M^{SJ} | x, n) = \frac{m(x|n, SJ)}{m(x|n, SJ) + m(x|n, GSJ)} \]
and
\[ \Pr(M^{GSJ}|x, n) = 1 - \Pr(M^{SJ}|x, n). \]

where \( m(x|n, w) = \int_0^1 \Pr(x|n, \theta, w)\text{Beta}(\theta|1/2, 1/2)d\theta, \text{ we } \{SJ, GSJ\}. \)

### 5.4 Testing the Equality of Treatment Effectiveness

Comparison of treatment effectiveness based on multiple studies is of great significance in meta-analysis. Suppose that two alternative treatments \( T_1 \) and \( T_2 \) are applied to patients in \( N_1 \) and \( N_2 \) healthcare centres and let \( \{x_s, n_s, s = 1, 2\} \) be the centre effectiveness samples of the treatments. Let \( \{\Pr(x_1|n_1, \alpha), \pi(\alpha)\} \) and \( \{\Pr(x_2|n_2, \beta), \pi(\beta)\} \) denote the conditional likelihoods and priors of the meta-effectiveness \( \alpha \) and \( \beta \) of the treatments, respectively.

Here we wish to test that effectiveness of two treatments is same. The null hypothesis proposed to be tested is

\( H_0 : \alpha = \beta \)

versus \( H_1 : (\alpha, \beta) \in (0, 1) \times (0, 1). \)

This is equivalent to the model selection problem between the Bayesian models

\( M_0 : \{\Pr(x_1|n_1, \theta)\Pr(x_2|n_2, \theta), \pi(\theta)\} \)

and

\( M_1 : \{\Pr(x_1|n_1, \alpha)\Pr(x_2|n_2, \beta), \pi(\alpha, \beta)\} \)

where it is assumed that \( x_1 \) and \( x_2 \) are independent, conditional on the meta-parameters.

The optimal solution to this decision problem under a 0-1 loss function is to reject the null \( H_0 \) if its posterior probability is smaller than \( 1/2 \), that is, if \( \Pr(M_0|x_1, n_1, x_2, n_2) < 1/2. \) For simplicity, we use the notation \( \Pr(M_0|x_1, x_2) \) instead of \( \Pr(M_0|x_1, n_1, x_2, n_2) \). Assuming that a priori \( \Pr(M_0) = \Pr(M_1) = 1/2 \), the posterior probability under \( M_0 \) is given by

\[ \Pr(M_0|x_1, x_2) = \frac{1}{1 + B_{10}(x_1, x_2)}, \]
where \( B_{10} \): Bayes factor for comparing \( M_1 \) and \( M_0 \) is given as

\[
B_{10}(x_1, x_2) = \frac{\int_0^1 \int_0^1 Pr(x_1|n_1, \alpha)Pr(x_2|n_2, \beta)\pi(\alpha, \beta)d\alpha d\beta}{\int_0^1 Pr(x_1|n_1, \theta)Pr(x_2|n_2, \theta)\pi(\theta)d\theta}.
\]

For uniform marginals,

Writing

\[
m_U^0(x_1, x_2|w_1, w_2) = \int_0^1 \left\{ \prod_{s=1}^2 Pr(x_s|n_s, \theta, w_s) \right\} d\theta, \quad w_1, w_2 \in \{FGM, GFGM\}
\]

\[
m_U^1(x_1, x_2|w_1, w_2) = \prod_{s=1}^2 \left\{ \int_0^1 Pr(x_s|n_s, \alpha, w_s) d\alpha \right\}, \quad w_1, w_2 \in \{FGM, GFGM\}.
\]

\[
m_U(x_1, x_2) = \sum_{w_1, w_2} \sum_{j=0}^1 m_U^j(x_1, x_2|w_1, w_2).
\]

The posterior probability of \( M_0 \) is given as

\[
Pr_U^U(M_0|x_1, x_2) = \frac{\sum_{w_1, w_2} m_U^0(x_1, x_2|w_1, w_2)}{m_U(x_1, x_2)}.
\]

We note that \( m_U^0(x_1, x_2|w_1, w_2) \) (\( m_U^1(x_1, x_2|w_1, w_2) \)) denote the marginal of the data under the null hypothesis when the distribution of \( x_1 \) comes from the link \( w_1 \) and that of \( x_2 \) from \( w_2 \).

\( Pr_U^U(M_0|x_1, x_2) \) can be written as a mixture of the posterior probability of the null under the links \( \{FGM, GFGM\} \) where the weights of the mixture are their posterior probabilities, that is,

If for \( w_1, w_2 \in \{FGM, GFGM\} \),

\[
Pr(M_0|x_1, x_2, w_1, w_2) = \frac{m_U^0(x_1, x_2|w_1, w_2)}{\sum_{j=0}^1 m_U^j(x_1, x_2|w_1, w_2)}
\]

and

\[
Pr(M_{w_1,w_2}|x_1, x_2) = \frac{\sum_{j=0}^1 m_U^j(x_1, x_2|w_1, w_2)}{m_U(x_1, x_2)}.
\]
then \( P_{r^U}(M_0|\mathbf{x}_1, \mathbf{x}_2) = \sum_{w_1, w_2} Pr(M_0|\mathbf{x}_1, \mathbf{x}_2, w_1, w_2)Pr(M^{w_1, w_2}|\mathbf{x}_1, \mathbf{x}_2). \)

Under the Jeffrey marginals and for \( w_1, w_2 \in \{SJ, GSJ\}, \) the posterior probability under \( M_0 \) is obtained as

\[
P_{r^J}(M_0|\mathbf{x}_1, \mathbf{x}_2) = \frac{\sum_{w_1, w_2} m_0^J(\mathbf{x}_1, \mathbf{x}_2|w_1, w_2)}{m^J(\mathbf{x}_1, \mathbf{x}_2)}
\]

where

\[
m_0^J(\mathbf{x}_1, \mathbf{x}_2|w_1, w_2) = \frac{1}{\pi} \int_0^1 \left\{ \prod_{s=1}^2 Pr(\mathbf{x}_s|\mathbf{n}_s, \theta, w_s) \right\} \theta^{-1/2}(1 - \theta)^{-1/2} d\theta,
\]

\[
m_1^J(\mathbf{x}_1, \mathbf{x}_2|w_1, w_2) = \prod_{s=1}^2 \left\{ \frac{1}{\pi} \int_0^1 Pr(\mathbf{x}_s|\mathbf{n}_s, \alpha, w_s)\alpha^{-1/2}(1 - \alpha)^{-1/2} d\alpha \right\},
\]

\[
m^J(\mathbf{x}_1, \mathbf{x}_2) = \sum_{w_1, w_2} \sum_{j=0}^1 m_j^J(\mathbf{x}_1, \mathbf{x}_2|w_1, w_2).
\]

### 5.5 Real Life Illustrations

For illustration purpose, the analysis of four real data sets is presented using the uniform and Jeffrey marginals.

#### 5.5.1 Example 1

A real data set from (Fisher and Van Belle (1983)) summarizing 15 independent clinical trials of progabide, an anti-epileptic drug is considered. The response variable is whether a greater than 50 in patient seizure frequency occurred, compared with a baseline in six open studies (indexed by 1) and compared with placebo for nine closed studies (indexed by 2). An open study is here a trial that permits the investigator to know the treatment regime and a closed study which is a double-blind controlled clinical trial. The data sets are

\[
\mathbf{n}_1 = [30, 16, 69, 23, 42, 151],
\]

\[
\mathbf{x}_1 = [17, 8, 41, 13, 32, 90],
\]

\[
\mathbf{n}_2 = [20, 20, 17, 15, 18, 17, 19, 51, 59],
\]

\[
\mathbf{x}_2 = [5, 9, 3, 7, 8, 9, 1, 12, 17].
\]
For these data, we test the null hypothesis \( H_{01} \) that the efficacy of progabide in open and closed studies is the same. The posterior probabilities of the FGM, GFGM, SJ and GSJ models are given in Table 5.1.

**Table 5.1: Posterior Model Probabilities**

<table>
<thead>
<tr>
<th>Uniform Marginals</th>
<th>Jeffrey Marginals</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Pr(M_{FGM</td>
<td>FGM}^{\text{FGM}}</td>
</tr>
<tr>
<td>2.795526 \times 10^{-4312}</td>
<td>2.795526 \times 10^{-4312}</td>
</tr>
<tr>
<td>( Pr(M_{GFGM</td>
<td>GFGM}^{\text{GFGM}}</td>
</tr>
<tr>
<td>3.570176 \times 10^{-407}</td>
<td>1.167761 \times 10^{-396}</td>
</tr>
</tbody>
</table>

The values in Table 5.1 depict that the data favour large homogeneity for uniform marginals and moderate heterogeneity for Jeffrey marginals. The values of the posterior probabilities of \( H_{01} \) under uniform and Jeffrey marginals are given in Table 5.2.

**Table 5.2: Posterior probability under \( H_{01} \)**

<table>
<thead>
<tr>
<th>Uniform Marginals</th>
<th>Jeffrey Marginals</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Pr^U(M_0</td>
<td>x_1, x_2) )</td>
</tr>
<tr>
<td>0.916543</td>
<td>0.416915</td>
</tr>
</tbody>
</table>

These values indicate a strong empirical evidence to accept the equality of efficacy of progabide for open and closed studies for uniform marginals. But null hypothesis of the equality of efficacy of progabide for open and closed studies evidence for the Jeffrey marginals is rejected.
5.5.2 Example 2

A data set analyzed by Bellamy et al. (2009) is considered where 20 studies were selected in order to assess the strength of association between type 2 diabetes mellitus and gestational diabetes. This association can have serious implications for glucose tolerance disorders in women, as well as for the prevention or delay of the development of type 2 diabetes. The total number of women with type 2 diabetes is 675,455 and 31,867 of these had previous pregnancies affected by gestational diabetes, with a total of 10,859 incident cases of type 2 diabetes. The finding was that women with gestational diabetes had an increased risk of developing type 2 diabetes as compared to those who had a normoglycemic pregnancy.

The data sets are

\[ n_1 = [21823, 620, 68, 166, 295, 5470, 70, 35, 23, 435, 696, 229, 28, 45, 801, 15, 241, 47, 615, 145], \]
\[ x_1 = [2874, 71, 21, 43, 53, 405, 6, 13, 7, 23, 44, 21, 10, 15, 105, 10, 33, 14, 224, 5], \]
\[ n_2 = [637341, 868, 39, 2242, 111, 783, 108, 489, 11, 435, 70, 61, 52, 39, 431, 35, 57, 47, 328, 41], \]
\[ x_2 = [6628, 22, 0, 150, 1, 16, 7, 8, 0, 0, 0, 1, 0, 1, 7, 0, 0, 3, 18, 0]. \]

where indices 1 and 2 refer to data sets with and without gestational diabetes respectively.

For these data sets, we test the null hypothesis \( H_{02} \) that the probability to develop type 2 diabetes for women with gestational diabetes is same as that for women with a normoglycemic pregnancy. For this purpose, the posterior model probabilities have been computed and these values are given in Table 5.3.
Table 5.3: Posterior Model Probabilities

<table>
<thead>
<tr>
<th>Uniform Marginals</th>
<th>Jeffrey Marginals</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Pr(M_{GPM,GPM}^{GPM}</td>
<td>x_1, x_2) )</td>
</tr>
<tr>
<td>( 5.103193 \times 10^{-18955} )</td>
<td>( 3.211399 \times 10^{-855} )</td>
</tr>
<tr>
<td>( Pr(M_{GPM,GP}^{GPM}</td>
<td>x_1, x_2) )</td>
</tr>
<tr>
<td>( 2.665447 \times 10^{-9108} )</td>
<td>1</td>
</tr>
</tbody>
</table>

Using the values in Table 5.3, the posterior probabilities of the null model \( H_{02} \) have been computed and shown in Table 5.4.

Table 5.4: Posterior probability under \( H_{02} \)

<table>
<thead>
<tr>
<th>Uniform Marginals</th>
<th>Jeffrey Marginals</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Pr^U(M_0</td>
<td>x_1, x_2) )</td>
</tr>
<tr>
<td>0.903502</td>
<td>0.852672</td>
</tr>
</tbody>
</table>

From the values of posterior probabilities in Table 5.4, it is concluded that under the uniform and the Jeffrey marginals, there is moderate empirical evidence for the equality of developing type 2 diabetes for the population of women with and without gestational diabetes.

5.5.3 Example 3

The data are from 16 randomised controlled trials of intravenous magnesium in the prevention of death following myocardial infarction (ISIS-4) (Collins et al. (1995); Eggar and Smith (1995); Egger et al. (1997)). The data sets are:
\( n_1 = [40, 135, 200, 48, 150, 59, 25, 22, 76, 27, 89, 23, 130, 1159, 107, 29011], \)
\( x_1 = [1, 9, 2, 1, 10, 1, 1, 0, 6, 1, 2, 5, 4, 90, 4, 2216], \)
\( n_2 = [36, 135, 200, 46, 148, 56, 23, 21, 75, 27, 80, 33, 122, 1157, 108, 29039], \)
\( x_2 = [2, 23, 7, 1, 8, 9, 3, 1, 11, 7, 12, 13, 8, 118, 17, 2103]. \)

where \( n_1 (n_2) \) is the number of patients under intervention (control) group,
\( x_1 (x_2) \) is the number of incidence of deaths for intervention (control) group.

For these data, the posterior probabilities of models FGM, GFGM, SJ and GSJ have been computed and displayed in Table 5.5.

<table>
<thead>
<tr>
<th>Uniform Marginals</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Pr(M^{FGM,FGM}</td>
</tr>
<tr>
<td>( 1.984610 \times 10^{-9655} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Jeffrey Marginals</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Pr(M^{SS}</td>
</tr>
<tr>
<td>( 0.999979 )</td>
</tr>
</tbody>
</table>

These posterior probabilities indicate that the data favour models with low heterogeneity for the uniform marginals and Jeffrey marginals.

The null hypothesis \( H_{03} \) to be tested is that the incidence of deaths for patients under intervention group and for patients who are under control group is the same. Using values in Table 5.5, the posterior probabilities are computed under \( H_{03} \). These values are displayed in Table 5.6 and they lead to the conclusion that there is moderate empirical evidence to accept that the incidence of deaths under intervention group and control group is the same both under the uniform and Jeffrey marginals. This evidence is slightly stronger when using the uniform marginals.
Table 5.6: Posterior probability under $H_{03}$

<table>
<thead>
<tr>
<th>Uniform Marginals</th>
<th>Jeffrey Marginals</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Pr^U(M_0</td>
<td>x_1, x_2)$</td>
</tr>
<tr>
<td>0.896242</td>
<td>0.848938</td>
</tr>
</tbody>
</table>

5.5.4 Example 4

Niel-Weise et al. (2007) performed a meta-analysis on the effect of anti-infective-treated central venous catheters on catheter-related bloodstream infection (CRBSI) in the acute care setting. The analysis comprised the data of 18 clinical trials comparing the risk of CRBSI in patients with an anti-infective-treated catheter and patients with a standard catheter. The data sets are:

$n_1 = [117, 35, 195, 136, 157, 139, 177, 39, 103, 122, 64, 58, 175, 180, 105, 262, 362, 69]$,

$x_1 = [3, 3, 9, 7, 6, 4, 3, 2, 19, 2, 7, 1, 5, 11, 0, 1, 3, 1]$.

$n_2 = [116, 44, 208, 130, 151, 98, 174, 74, 97, 113, 66, 70, 188, 187, 118, 252, 345, 64]$,

$x_2 = [0, 1, 2, 0, 5, 1, 1, 1, 1, 0, 0, 3, 6, 0, 0, 1, 4]$.

where $n_1$ ($n_2$): the number of patients under Standard (Anti-infective) catheter,

$x_1$ ($x_2$): the number of incidence of CRBSIs.

The null hypothesis $H_{04}$ to be tested is that the the incidence of CRBSIs for patients under Standard catheter and for patients who are under Anti-infective catheter is the same.

The posterior model probabilities are shown in Table 5.7.
Table 5.7: Posterior Model Probabilities

<table>
<thead>
<tr>
<th>Uniform Marginals</th>
<th>Jeffrey Marginals</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Pr(M_{FGM,FGM}^{</td>
<td></td>
</tr>
<tr>
<td>$2.451099 \times 10^{-4354}$</td>
<td>$6.489629 \times 10^{-4392}$</td>
</tr>
<tr>
<td>$1.000$</td>
<td>$1.000$</td>
</tr>
</tbody>
</table>

These posterior probabilities indicate that the data favour models with low heterogeneity for the uniform as well as Jeffrey marginals.

Using values in Table 5.7, the posterior probabilities of the null model under uniform and Jeffrey marginals are given in Table 5.8.

Table 5.8: Posterior probability under $H_{04}$

<table>
<thead>
<tr>
<th>Uniform Marginals</th>
<th>Jeffrey Marginals</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Pr^U(M_0</td>
<td>x_1, x_2)$</td>
</tr>
<tr>
<td>$0.903502$</td>
<td>$0.912095$</td>
</tr>
</tbody>
</table>

These values lead to the conclusion that there is strong empirical evidence to accept that the incidence of CRBSIs under Standard catheter and Anti-infective catheter is same under the uniform as well as Jeffrey marginals.
5.6 Conclusions

We consider classes of bivariate prior distributions for \((\theta_i, \theta)\) such that they have given marginals and can be parameterized in terms of the Pearson’s correlation coefficient between \(\theta_i\) and \(\theta\). The choice of the linking distributions is driven by the linear correlation coefficient. For these classes, it is seen that higher the correlation coefficient, the smaller is the between-center heterogeneity.

We have used the FGM (Farlie-Gumbel-Morgenstern) and GFGM (Generalized Farlie-Gumbel-Morgenstern) classes of distributions when assuming uniform marginals and the SJ (Sarmanov Jeffreys) and GSJ (Generalized Sarmanov Jeffreys) classes in case of Jeffrey marginals. Bayesian model averaging has been used for pooling the inference conditional on models FGM and GFGM and SJ and GSJ models.

A Bayesian procedure for testing the equality of the meta-parameter of two alternative treatments have been proposed. The final testing result is obtained as a mixture of the tests for each of the links. The weights of the mixture are the posterior probabilities of the considered models which inform us about the linking distribution favoured by the data.

A particular dependence structure between \(\theta_i\) and \(\theta\) is considered that model a wide range of heterogeneity degrees.