Journal of Data Science and Modeling, Vol. 2, No. 2, 15-34, June 2024

Research Manuscript

# Bayesian Semiparametric Meta-Regression Model

Ehsan Ormoz*[1], Farzad Eskandari[2]

[1]Department of Mathematics and Statistics, Mashhad Branch, Islamic Azad University, Mashhad, Iran.

[2]Department of Statistics, Allameh Tabataba'i University, Tehran, Iran.

Recieved: 25/10/2024 Accepted: 10/02/2025

**Abstract:**

This paper introduces a novel semiparametric Bayesian approach for bivariate meta-regression. The method extends traditional binomial models to trinomial distributions, accounting for positive, neutral, and negative treatment effects. Using a conditional Dirichlet process, we develop a model to compare treatment and control groups across multiple clinical centers. This approach addresses the challenges posed by confounding factors in such studies. The primary objective is to assess treatment efficacy by modeling response outcomes as trinomial distributions. We employ Gibbs sampling and the Metropolis-Hastings algorithm for posterior computation. These methods generate estimates of treatment effects while incorporating auxiliary variables that may influence outcomes. Simulations across various scenarios demonstrate the model's effectiveness. We also establish credible intervals to evaluate hypotheses related to treatment effects. Furthermore, we apply the methodology to real-world data on economic activity in Iran from 2009 to 2021. This application highlights the practical utility of our approach in meta-analytic contexts. Our research contributes to the growing body of literature on Bayesian methods in meta-analysis. It provides valuable insights for improving clinical study evaluations.

**Keywords:** Bayesian Model Selection, Bayesian Semi-parametric, Meta-Analysis, Meta-Regression, Multinomial Distribution.

**Mathematics Subject Classification (2010):** 33C45, 49-XX.

*Corresponding Author: ehsanormoz@mshdiau.ac.ir

# 1. Introduction

In the context of meta-analysis, we often encounter situations where $m$ centers are considered for conducting a similar clinical study to compare a treatment with a control group. Early work in meta-analysis includes combining effect size estimates or merging $p$-values (Tippett , 1931; Pearson , 1933).

Burr and Doss (2005) describe a Bayesian semiparametric analysis for clinical studies, focusing on cases where $p_i^{(1)}$ and $p_i^{(2)}$ represent the success probabilities of two distinct treatment and control groups in center $i$. These probabilities are compared to evaluate improvement. For $j = 1, 2$, let $y_i^{(j)}$ represent the binomial outcome in center $i$. Thus, for $i = 1, 2, \ldots, m$, we have:

$$y_i^{(j)} \sim \text{Binomial}(n_i^{(j)}, p_i^{(j)}), \qquad\qquad j = 1, 2 \tag{1.1}$$

Burr and Doss (2005) extend the traditional binomial model to handle trinomial outcomes. This framework is particularly useful in real-world applications where outcomes are more nuanced. For example, in evaluating a proposed drug, three outcomes can be considered: the drug improves the patient's symptoms, worsens them, or has no effect. To account for these outcomes and their interdependence, the normal prior in their work is extended to a bivariate normal distribution. This extension enables the model to incorporate correlations between outcomes, providing a more comprehensive analysis.

A parametric Bayesian approach in meta-analysis has been developed by various authors, including Carlin (1992), Higgins (1997), and Maier et al. (2022). However, in recent years, considerable attention has been directed toward nonparametric and semiparametric Bayesian approaches. Chung and Dunson (2007) attribute this trend to the efficiency and simplicity of posterior computation in Dirichlet process mixtures. Related approaches have examined semiparametric models in meta-analysis, including Burr and Doss (2005), Ohlsen et al. (2007), and Frömke et al. (2022). Dominici and Parmigiani (2001) and Carota and Parmigiani (2002) also focused on semiparametric Bayesian approaches for count data in a distinct framework.

However, substantial heterogeneity is often observed among studies, and it is the task of statisticians to assess potential sources of this heterogeneity (Thompson , 1994). In the context of meta-analysis, auxiliary study-level variables can be employed to explain differences between studies. The term meta-regression, used to describe such an analysis, dates back to works by Bashore et al. (1989), Jones (1992), and Greenland (1994).

Thompson (1994) argues that heterogeneity can be regarded as a valuable tool, as it enables the application of beneficial approaches that aim to examine the impact of potential sources of heterogeneity on the overall treatment effect. For instance, the treatment effect might be lower in studies involving a higher number of older men compared to those with more young women. The dependence of treatment effects on one or more characteristics, such as age and gender, can be examined using meta-regression.

In meta-regression, subject characteristics are considered as auxiliary variables in a regression analysis to estimate treatment effects. As stated by Armitage and Colton (1998), to reduce post-study risk due to examining existing data, such auxiliary variables should be pre-specified before starting the study. The statistical goal of meta-regression is to explain the variance component among subjects using auxiliary variables.

Two main ideas are considered in this paper. First, we focus on the method used to create a class $\Gamma$ close to the specified prior $\pi_0$ with a domain close to $\Omega$. The second idea is the computational technique used to calculate the posterior distribution. Specifically, we aim to extend the results of Burr and Doss (2005) to a bivariate meta-regression.

In Section 2, we introduce the proposed model based on the conditional Dirichlet process. Posterior distributions are calculated in Section 3. Since these posterior distributions lack closed-form expressions, they are approximated using simulation techniques. To estimate parameters in a Bayesian manner, we employ Gibbs sampling and the Metropolis-Hastings algorithm, discussed in Section 4. In Section 5, the effectiveness of the proposed method is evaluated using three simulated datasets, and a practical example is provided in Section 6.

## 2. Model Description

In clinical studies, situations often arise where, in addition to the positive and negative effects of a drug or treatment, a third neutral effect state is also relevant. In such cases, the response variable for the treatment effect follows a trinomial distribution. Here, we aim to generalize studies conducted on the binomial distribution to a trinomial case using a semiparametric Bayesian approach within the meta-regression framework.

Assume a similar study is conducted in $m$ centers, where in each center, a comparison between the treatment and control groups is made, yielding outcomes labeled as positive, neutral, and negative effects. Therefore, the response variable for both the treatment and control groups follows a trinomial distribution. Consequently, for patient group $j$ in center $i$, we have:

$$r_i^{(j)} = (r_i^{(j1)}, r_i^{(j2)}, r_i^{(j3)}) \sim \text{Trinomial}(n_i^{(j)}, \mathbf{p}_i^{(j)}), \qquad i = 1, \ldots, m; \qquad j = 1, 2 \tag{2.2}$$

where the frequency of the $l$th outcome, $r_i^{(jl)}, l = 1,2,3$, satisfies the conditions:

$$0 \leq r_i^{(jl)} \leq n_i^{(j)} \quad \text{and} \quad \sum_{l=1}^{3} r_i^{(jl)} = n_i^{(j)}.$$

Furthermore, the corresponding probability vector $\mathbf{p}_i^{(j)}$ satisfies:

$$0 < p_i^{(jl)} < 1 \quad \text{and} \quad \sum_{l=1}^{3} p_i^{(jl)} = 1.$$

In such studies, the primary goal is to compare the treatment and control groups by testing the following hypothesis:

$$H_0 : \mathbf{p}_i^{(1)} = \mathbf{p}_i^{(2)}, \qquad i = 1, \ldots, m. \tag{2.3}$$

These comparisons are often confounded by unavoidable effects due to differences between the testing centers. Thus, to test $H_0$, these confounding factors must be accounted for. This issue is central to many meta-analysis studies. In this paper, we address it by regressing $\mathbf{p}_i^{(j)}$ on some auxiliary variables that specify the conditions under which the experiment was conducted. After obtaining the regression model, we can make inferences about $\mathbf{p}_i^{(j)}$ while accounting for the effects of confounding factors. In this approach, we use the log-odds of the treatment and control outcomes, denoted by $D_i = \left(D_i^{(1)}, D_i^{(2)}\right)$, where

$$D_i^{(k)} = \log \frac{p_i^{(1k)}}{p_i^{(2k)}} = x_i'^{(k)} \beta_i^{(k)}, \qquad k = 1, 2. \tag{2.4}$$

Note that here $x_i'^{(k)}$ represents a vector of auxiliary variables. It is evident that equations 2.3 and 2.4 pursue similar objectives. To perform Bayesian analysis, we require the prior distribution for $\beta_i$. We assume these prior distributions are bivariate normal with means dependent on auxiliary variables representing experimental conditions, as follows:

$$\beta_i = (\beta_i^{(1)}, \beta_i^{(2)})' \sim^{\text{ind}} N_2(\boldsymbol{\eta}_i, \Sigma_i),$$

where

$$\boldsymbol{\eta}_i = (\eta_i^{(1)}, \eta_i^{(2)})', \qquad \Sigma_i = \begin{pmatrix} \sigma_i^{2(1)} & \sigma_i^{(12)} \\ \sigma_i^{(21)} & \sigma_i^{2(2)} \end{pmatrix}.$$

Thus,

$$D_i = (D_i^{(1)}, D_i^{(2)})' = (x_i'^{(1)} \beta_i^{(1)}, x_i'^{(2)} \beta_i^{(2)})' \sim^{\text{ind}} N_2(\boldsymbol{\eta}_i^*, \Sigma_i^*), \tag{2.5}$$

where

$$\boldsymbol{\eta}_i^* = \begin{pmatrix} \eta_i^{(1)*} \\ \eta_i^{(2)*} \end{pmatrix} = (x_i'^{(1)} \eta_i^{(1)}, x_i'^{(2)} \eta_i^{(2)})', \quad \Sigma_i^* = \begin{pmatrix} x_i'^{(1)} & x_i'^{(2)} \end{pmatrix} \Sigma_i \begin{pmatrix} x_i^{(1)} \\ x_i^{(2)} \end{pmatrix}.$$

Note that $\eta_i^{(k)}$ are the regression coefficients, and $\rho_i = \frac{\sigma_i^{(12)}}{\sqrt{\sigma_i^{2(1)}\sigma_i^{2(2)}}}$ is the correlation coefficient between $D_i^{(1)}$ and $D_i^{(2)}$, which arises from the common experimental environment for treatment and control groups. By considering the prior distribution for the vector $\boldsymbol{\eta}_i$ and transforming from $\boldsymbol{\eta}_i$ to $\mathbf{p}$, we can obtain the prior distribution for $\mathbf{p}$. It should be noted that although we use the specified model corresponding to equations 2.4 and 2.5, these settings can be generalized to more complex scenarios, leading to a bivariate meta-regression problem.

As Griffin and Steel (2007) mentioned, the Dirichlet process (Ferguson , 1973) has been extensively used as a prior distribution for an unknown model distribution, especially when dealing with multinomial distributions. In the following, we will specify the required prior distributions using the Dirichlet process, and consequently estimate posterior distributions and their parameters under a Bayesian semiparametric meta-regression model.

**Definition 2.1.** *Let $\Theta$ be a set and $A$ be the sigma-field of subsets of $\Theta$. Also, let $\alpha$ be a finite, non-empty, non-negative, and finitely additive measure on $(\Theta, A)$. We say that a random probability measure $P$ on $(\Theta, A)$ is a Dirichlet process with parameter $\alpha$ on $(\Theta, A)$, and we write $P \sim D(\alpha)$, if for every $k = 1, 2, \ldots,$ and for every measurable partition $B_1, \ldots, B_k$ of $\Theta$, the joint distribution of the random probabilities $(P(B_1), \ldots, P(B_k))$ is Dirichlet with parameters $(\alpha(B_1), \ldots, \alpha(B_k))$.*

**Definition 2.2.** *Let $H_\theta$ for $\theta \in \Theta \subset \Re^k$ be a parametric family of distributions on the real line, and let $\lambda$ be a distribution on $\Theta$. Also, let $M_\theta > 0$ be known weights, and set $\alpha_\theta = M_\theta H_\theta$. If $\theta$ is sampled from $\lambda$ and $F$ is sampled from $D_{\alpha_\theta}$, i.e., a Dirichlet process with parameter $\alpha_\theta$, then we say that the prior distribution for $F$ is a mixture of Dirichlet processes (Antoniak , 1974).*

In other words, a mixture of Dirichlet processes is a Dirichlet process whose measure parameter is itself a random variable. For simplicity, we assume that $M$ does not depend on $\theta$. Doss (1985) defined conditional Dirichlet processes as follows.

**Definition 2.3.** *Let $\alpha$ be a finite measure on the real line, and $\mu \in (-\infty, \infty)$ be a fixed value. Instead of an arbitrary set $A$, let $\alpha_-^\mu$ and $\alpha_+^\mu$ be the restrictions of $\alpha$ to $(-\infty, \mu)$ and $(\mu, \infty)$, respectively, defined as:*

$$\alpha_-^\mu(A) = \alpha(A \cap (-\infty, \mu)) + \frac{1}{2}\alpha(A \cap \{\mu\}),$$

$$\alpha_+^\mu(A) = \alpha(A \cap (\mu, \infty)) + \frac{1}{2}\alpha(A \cap \{\mu\}).$$

*$F_- \sim D_{\alpha_-^\mu}$ and $F_+ \sim D_{\alpha_+^\mu}$ are sampled independently, and $F(t)$ is defined as:*

$$F(t) = \frac{1}{2}F_-(t) + \frac{1}{2}F_+(t), \tag{2.6}$$

*where the distribution of $F$ is denoted by $D^{\mu}_{\alpha}$. Note that the median of $F$ is equal to $\mu$ with probability one. Therefore, if $F \sim D_{\alpha}$, then $D^{\mu}_{\alpha}$ is the conditional distribution of $F$ given that its median is equal to $\mu$.*

Similar to Burr and Doss (2005) and using equation 2.5, for $i = 1, 2, \ldots, m$ and $j = 1, 2$, we propose the following model:

$$\eta_i^{(j)} \mid F^{(j)} \sim^{\text{ind}} F^{(j)}; \tag{2.7}$$

$$F^{(j)} \mid \mu^{(j)}, \tau^{(j)} \sim^{\text{ind}} D^{\mu^{(j)}}_{M \times N(\mu^{(j)}, \tau^{2(j)})}; \tag{2.8}$$

$$\mu^{(j)} \mid \tau^{(j)} \sim^{\text{ind}} N(c^{(j)}, d^{(j)} \tau^{2(j)}); \tag{2.9}$$

$$\gamma^{(j)} = \frac{1}{\tau^{2(j)}} \sim^{\text{ind}} \Gamma(a^{(j)}, b^{(j)}). \tag{2.10}$$

**Remark 2.4.** *Note that, similar to Burr and Doss (2005), the use of an index for a distribution implies conditioning.*

The main question now is whether the means of $F^{(j)}$ differ significantly from zero. To answer this question, we will use posterior credible intervals.

## 3. Posterior Computation

Following a similar process to Burr and Doss (2005), the prior distribution of $\eta_i^{(j)}$ can be expressed as:

$$\pi_{\{\eta^{(j)}_{(-i)}, \mu^{(j)}, \tau^{(j)}\}}(\eta_i^{(j)}) = \frac{1}{2} \frac{M \times N_{-}^{\mu^{(j)}}(\mu^{(j)}, \tau^{2(j)}) + \sum_{k \neq i, \eta_k^{(j)} < \mu^{(j)}} \delta_{\eta_k^{(j)}}}{\frac{M}{2} + m_{-}^{(j)}} + \frac{1}{2} \frac{M \times N_{+}^{\mu^{(j)}}(\mu^{(j)}, \tau^{2(j)}) + \sum_{k \neq i, \eta_k^{(j)} > \mu^{(j)}} \delta_{\eta_k^{(j)}}}{\frac{M}{2} + m_{+}^{(j)}}, \tag{3.11}$$

where

$$m_{-}^{(j)} = \sum_{k \neq i} I(\eta_k^{(j)} < \mu^{(j)}) \quad \text{and} \quad m_{+}^{(j)} = \sum_{k \neq i} I(\eta_k^{(j)} > \mu^{(j)}). \tag{3.12}$$

This distribution is a special case of the normal distribution with hyperparameters $\boldsymbol{\alpha} = (\mu^{(1)}, \mu^{(2)}, \tau^{(1)}, \tau^{(2)})$. Using equation 2.4, the likelihood of $D_i$ conditional on $\eta_i^{(1)}, \eta_i^{(2)}$, $\sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i$ is given by:

$$L_{\{\eta^{(1)}_{(-i)}, \eta^{(2)}_{(-i)}, \boldsymbol{\alpha}\}}(D_i | \eta_i^{(1)}, \eta_i^{(2)}) = N_2(x_i'^{(1)} \eta_i^{(1)}, x_i'^{(2)} \eta_i^{(2)}, \sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i). \tag{3.13}$$

Thus, we aim to derive the posterior distribution

$$\pi_D(\eta_i^{(1)}, \eta_i^{(2)} | \eta^{(1)}_{(-i)}, \eta^{(2)}_{(-i)}, \mu^{(1)}, \mu^{(2)}, \tau^{(1)}, \tau^{(2)}),$$

where $\eta^{(j)}_{(-i)} = (\eta^{(j)}_1, \ldots, \eta^{(j)}_{i-1}, \eta^{(j)}_{i+1}, \ldots, \eta^{(j)}_m)$.

The following theorem demonstrates that while the posterior distribution is complex, it provides a better estimate of the parameters in equation 2.4.

**Theorem 3.1.** *Given equations 3.11 and 3.13, the posterior distribution of* $(\eta_i^{(1)}, \eta_i^{(2)})$ *conditional on* $(\eta^{(1)}_{(-i)}, \eta^{(2)}_{(-i)})$, $(\mu^{(1)}, \mu^{(2)})$, *and* $(\tau^{(1)}, \tau^{(2)})$ *is:*

$$\begin{aligned}\pi_D(\eta_i^{(1)}, \eta_i^{(2)} | \eta^{(1)}_{(-i)}, \eta^{(2)}_{(-i)}, \boldsymbol{\alpha}) \propto & W_0 + M \sum_{k \neq i} W_1^{ki} N_2(x_i'^{(1)} \eta_i^{(1)}, x_k'^{(2)} \eta_k^{(2)}, \sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i) \\ & + M \sum_{k \neq i} W_2^{ik} N_2(x_i'^{(1)} \eta_i^{(1)}, x_k'^{(2)} \eta_k^{(2)}, \sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i) \\ & + \sum_{k \neq i} \sum_{h \neq i} W_3^{kh} N_2(x_i'^{(1)} \eta_i^{(1)}, x_k'^{(2)} \eta_k^{(2)}, \sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i),\end{aligned}$$

*where for* $a \in \{-, +\}$ *and* $b \in \{-, +\}$*:*

$$\begin{aligned}S^1_{ki} &= \{\eta_k^{(1)}, \eta_i^{(2)} | \eta_k^{(1)} \in (-\infty, \mu^{(1)}), \eta_i^{(2)} \in (-\infty, \mu^{(2)})\}, \\ S^2_{ki} &= \{\eta_k^{(1)}, \eta_i^{(2)} | \eta_k^{(1)} \in (-\infty, \mu^{(1)}), \eta_i^{(2)} \in (\mu^{(2)}, +\infty)\}, \\ S^3_{ki} &= \{\eta_k^{(1)}, \eta_i^{(2)} | \eta_k^{(1)} \in (\mu^{(1)}, +\infty), \eta_i^{(2)} \in (-\infty, \mu^{(2)})\}, \\ S^4_{ki} &= \{\eta_k^{(1)}, \eta_i^{(2)} | \eta_k^{(1)} \in (\mu^{(1)}, +\infty), \eta_i^{(2)} \in (\mu^{(2)}, +\infty)\},\end{aligned}$$

*with*

$$C_{a,b} = \frac{M^2}{4} + \frac{M}{2}(m_a^{(1)} + m_b^{(2)}) + m_a^{(1)} m_b^{(2)},$$

*and*

$$\begin{aligned}W_0 &= \sum_{l=1}^{4} \sum_{a,b} C_{a,b}^{(1,2)} N_{a,b}^{(\mu^{(1)}, \mu^{(2)})}(A_i^{(1)}, A_i^{(2)}, B_i^{2(1)}, B_i^{2(2)}, \rho_i) I_{S^l_{ki}}(\eta_i^{(1)}, \eta_i^{(2)}), \\ W_1^{ik} &= \sum_{l=1}^{4} \sum_{a,b} \frac{1}{C_{a,b}} N_a^{\mu^{(1)}}(\mu^{(1)}, \tau^{2(1)}) I_{S^l_{ki}}(\eta_i^{(1)}, \eta_i^{(2)}), \\ W_2^{ki} &= \sum_{l=1}^{4} \sum_{a,b} \frac{1}{C_{a,b}} N_b^{\mu^{(2)}}(\mu^{(2)}, \tau^{2(2)}) I_{S^l_{ki}}(\eta_i^{(1)}, \eta_i^{(2)}), \\ W_3^{kh} &= \sum_{l=1}^{4} \sum_{a,b} \frac{1}{C_{a,b}} I_{S^l_{ki}}(\eta_i^{(1)}, \eta_i^{(2)}),\end{aligned}$$

*where for* $l = 1, 2$*:*

$$A_i^{(l)} = \frac{(1-\rho_i^2)\sigma_i^{2(l)} \mu^{(l)} + \tau^{2(l)} x_i'^{(l)} \beta_i^{(l)} x_i^{(l)}}{(1-\rho_i^2)\sigma_i^{2(l)} + x_i'^{(l)} \tau^{2(l)} x_i^{(l)}}, \quad B_i^{2(l)} = \frac{\sigma_i^{2(l)} \tau^{2(l)}}{(1-\rho_i^2)\sigma_i^{2(l)} + x_i'^{(l)} \tau^{2(l)} x_i^{(l)}}. \tag{3.14}$$

*Additionally,*

$$C_{a,b}^{(1,2)} = \frac{M^2}{C_{a,b}} N(x_i'^{(1)}\mu^{(1)}, (1-\rho_i^2)\sigma_i^{2(1)} + x_i'^{(1)}\tau^{2(1)}x_i^{(1)}) \times N(x_i'^{(2)}\mu^{(2)}, (1-\rho_i^2)\sigma_i^{2(2)} + x_i'^{(2)}\tau^{2(2)}x_i^{(2)}) e^{R_{a,b}}. \tag{3.15}$$

*Moreover,*

$$R_{a,b} = R \times U_{a,b},$$

*and*

$$R = -\frac{\rho_i}{1-\rho_i^2}\left\{\frac{1}{B_i^{(1)}B_i^{(2)}}(\eta_i^{(1)} - A_i^{(1)})(\eta_i^{(2)} - A_i^{(2)}) - \left(\frac{x_i'^{(1)}\beta_i^{(1)} - \eta_i^{(1)}}{\sigma_i^{(1)}}\right)\left(\frac{x_i'^{(2)}\beta_i^{(2)} - \eta_i^{(2)}}{\sigma_i^{(2)}}\right)\right\},$$

*with*

$$U_{-,-} = \begin{cases} 1 & (\eta_i^{(1)}, \eta_i^{(2)}) \in S_{ii}^1, \\ 0 & o.w. \end{cases}, \quad U_{-,+} = \begin{cases} 1 & (\eta_i^{(1)}, \eta_i^{(2)}) \in S_{ii}^2, \\ 0 & o.w. \end{cases},$$

$$U_{+,-} = \begin{cases} 1 & (\eta_i^{(1)}, \eta_i^{(2)}) \in S_{ii}^3, \\ 0 & o.w. \end{cases}, \quad U_{+,+} = \begin{cases} 1 & (\eta_i^{(1)}, \eta_i^{(2)}) \in S_{ii}^4, \\ 0 & o.w. \end{cases}.$$

Note that to generate $\boldsymbol{\alpha}$ from $\pi_D(\boldsymbol{\alpha}|\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)})$ based on equations 2.9 and 2.10, we have:

$$\pi_D(\boldsymbol{\alpha}|\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}) = \pi(\boldsymbol{\alpha}|\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}) = \pi(\mu^{(1)}, \tau^{2(1)}|\boldsymbol{\eta}^{(1)})\pi(\mu^{(2)}, \tau^{2(2)}|\boldsymbol{\eta}^{(2)}).$$

In this specific case, we use Proposition 1 from Burr and Doss (2005).

**Proposition 3.2.** *Burr and Doss (2005) Suppose $H$ is absolutely continuous with a continuous density function $h$ and its median is zero. Also, let $\psi_1, \dots, \psi_m$ be random samples from the distribution $F$, and the prior distribution for $F$ is a mixture of conditional Dirichlet processes $\int D^{\mu}_{M_\theta H_\theta}\lambda(d\theta)$. Then the posterior distribution $\theta$ conditional on $\psi_1, \dots, \psi_m$ is absolutely continuous with respect to $\lambda$, and*

$$\lambda_\psi(d\theta) = c(\psi)\left(\prod^{dist} h\left(\frac{\psi_i - \mu}{\tau}\right)\right) k(\psi, \theta) \times \left[\frac{(M_\theta)^{\#(\psi)}\Gamma(M_\theta)}{\Gamma(M_\theta + n)}\right]\lambda(d\theta),$$

*where*

$$K(\psi^{(j)}, \mu^{(j)}) = \left[\Gamma\left(\frac{M}{2} + \sum_{i=1}^{m} I(\psi_i^{(j)} < \mu^{(j)})\right) \times \Gamma\left(\frac{M}{2} + \sum_{i=1}^{m} I(\psi_i^{(j)} > \mu^{(j)})\right)\right]^{-1}. \tag{3.16}$$

In this equation, the symbol *dist* in the product operator indicates that the product is taken only over distinct values, $\#(\psi)$ is the number of distinct values in the vector $\psi$, and $c(\psi)$ is a normalization constant.

Now, using this proposition, we ensure that the posterior distributions have closed forms. Note that since the expression $K(\eta^{(j)}, \mu^{(j)})$ is analytically complex, in applications, simulation-based methods such as Gibbs sampling must be used.

**Theorem 3.3.** *Suppose $m^{(j)^*}$ is the number of distinct values of $\eta_i^{(j)}$ and define:*

$$\overline{\boldsymbol{\eta}^{(j)}} = \frac{1}{m^{(j)^*}} \sum^{dist} \eta_i^{(j)}.$$

*Then, $\pi(\mu^{(j)}, \tau^{(j)}|\boldsymbol{\eta}^{(j)})$ has a distribution proportional to the product:*

$$\pi(\mu^{(j)}, \tau^{(j)}|\boldsymbol{\eta}^{(j)}) = g_{\boldsymbol{\eta}^{(j)}}(\mu^{(j)}, \tau^{(j)}) K(\boldsymbol{\eta}^{(j)}, \mu^{(j)}), \tag{3.17}$$

*where $g_{\boldsymbol{\eta}^{(j)}}(\mu^{(j)}, \tau^{(j)})$ has a similar form to equations 2.9 and 2.10 with updated parameters $a^{(j)'}$, $b^{(j)'}$, $c^{(j)'}$, and $d^{(j)'}$, given by:*

$$\begin{aligned}
a^{(j)'} &= a^{(j)} + \frac{m^{(j)^*}}{2}, \\
b^{(j)'} &= b^{(j)} + \frac{1}{2}\sum^{dist}(\eta_i^{(j)} - \overline{\boldsymbol{\eta}^{(j)}})^2 + \frac{m^{(j)^*}(\overline{\boldsymbol{\eta}^{(j)}} - c^{(j)})^2}{2(1 + m^{(j)^*} d^{(j)})}, \\
c^{(j)'} &= \frac{c^{(j)} + m^{(j)^*} d^{(j)} \overline{\boldsymbol{\eta}^{(j)}}}{1 + m^{(j)^*} d^{(j)}}, \\
d^{(j)'} &= \frac{1}{m^{(j)^*} + d^{(j)^{-1}}}.
\end{aligned}$$

*$K(\psi^{(j)}, \mu^{(j)})$ is introduced in equation 3.16.*

**Remark 3.4.** *By integrating equation 3.17 with respect to $\tau$, we can derive the marginal distribution of $\mu^{(j)}$ conditional on $\boldsymbol{\eta}^{(j)}$, which is proportional to:*

$$t\left(2a^{(j)'}, c^{(j)'}, b^{(j)'} d^{(j)'}/a^{(j)'}\right)(.) K\left(\boldsymbol{\eta}^{(j)}, .\right), \tag{3.18}$$

*where $t(d, l, s^2)$ is a t-distribution with $d$ degrees of freedom, location parameter $l$, and scale parameter $s$. Moreover, conditional on $\mu^{(j)}$ and $\boldsymbol{\eta}^{(j)}$, the conditional distribution of $\frac{1}{\tau^{(j)}}$ is given by:*

$$\Gamma\left(a^{(j)'} + \frac{1}{2}, b^{(j)'} + \frac{(\mu^{(j)} - c^{(j)'})^2}{2d^{(j)'}}\right). \tag{3.19}$$

# 4. Gibbs Sampler

The proposed Gibbs sampler consists of two main steps:

**Step 1:** Updating $(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)})$. For $i = 1, \ldots, m$, generate the values of $(\eta_i^{(1)}, \eta_i^{(2)})$ conditional on the current values of $(\eta_j^{(1)}, \eta_j^{(2)})$ for $j \neq i$, $(\mu^{(1)}, \mu^{(2)})$, $(\tau^{(1)}, \tau^{(2)})$, and the data, iteratively.

**Step 2:** Updating $(\mu^{(1)}, \mu^{(2)}, \tau^{(1)}, \tau^{(2)})$. To generate $(\mu^{(1)}, \mu^{(2)}, \tau^{(1)}, \tau^{(2)})$ conditional on $(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)})$, the following two steps are performed:

**Step 2-a:** Generate $(\mu^{(1)}, \mu^{(2)})$ from the marginal distribution conditional on $(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)})$ as derived in equation 3.18. This distribution is proportional to the product of two $t$-distributions as in equation 3.18, multiplied by a factor that can be easily computed.

**Step 2-b:** Generate $(\tau^{(1)}, \tau^{(2)})$ conditional on $(\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)})$ and $(\mu^{(1)}, \mu^{(2)})$ from the distribution given in equation 3.19. For $j = 1, 2$, the distribution of $\frac{1}{\tau^{(j)}}$ follows a Gamma distribution, and $\tau^{(1)}$ and $\tau^{(2)}$ are independent.

## 5. Simulation

To evaluate the proposed model, we conducted a simulation study. Three scenarios were considered. In the first scenario, each center consists of two trinomial populations with probability vectors $\mathbf{p} = (0.5, 0.4, 0.1)$. In the second scenario, the two trinomial populations have different probability vectors, with the first population having $\mathbf{p} = (0.6, 0.3, 0.1)$ and the second having $\mathbf{p} = (0.5, 0.3, 0.2)$. Finally, in the third scenario, the probabilities of the two populations are $\mathbf{p} = (0.5, 0.4, 0.1)$ and $\mathbf{p} = (0.3, 0.6, 0.1)$, respectively. For each population, only one auxiliary variable was considered, and both auxiliary variables were generated from the $N(2, 1)$ distribution.

In each case, ten centers were used, and observations for each center were generated based on the described scenarios. Using these generated data, we can now evaluate the proposed model. To do so, we first compute the log-odds and their correlations. The variances were estimated using the following relation:

$$\mathrm{var}(\hat{\theta}) = \frac{F_1}{n_1 S_1} + \frac{F_2}{n_2 S_2},$$

where $\hat{\theta}$ is the observed log-relative risk for $i = 1, 2$, $S_i$ is the number of successes in the first cell of population $i$, $F_i$ is the number of failures, and $n_i$ is the number of observations for population $i$. The correlation was set to $-0.5$.

Using this data, we implemented the model. The estimates, mean squared errors (MSE), and confidence intervals for $\beta^{(1)}, \beta^{(2)}, D^{(1)}, D^{(2)}, \mu^{(1)}$, and $\mu^{(2)}$ are

Table 1: *Estimates, Mean Squared Errors, and Confidence Intervals for* $\beta^{(1)}$ *and* $\beta^{(2)}$

| | | $\beta^{(1)}$ | $\beta^{(2)}$ |
|---|---|---|---|
| Estimate | Scenario 1 | 0.0026 | −0.0235 |
| | Scenario 2 | 0.0681 | −1.1924 |
| | Scenario 3 | 0.5308 | −0.5094 |
| MSE | Scenario 1 | $7.17 \times 10^{-6}$ | 0.0005 |
| | Scenario 2 | 0.0046 | 1.4240 |
| | Scenario 3 | 0.2818 | 0.2598 |
| Confidence Interval | Scenario 1 | $(-0.0265, 0.0336)$ | $(-0.0660, 0.0060)$ |
| | Scenario 2 | $(0.0340, 0.1018)$ | $(-4.5655, 0.0558)$ |
| | Scenario 3 | $(0.1839, 0.7473)$ | $(-1.0539, -0.0057)$ |

reported in Tables 1-3. To compute these quantities, the Gibbs sampler was run for 1500 iterations, with the first 500 iterations discarded as burn-in. (It should be noted that convergence was checked using various criteria, and convergence was confirmed.)

As seen, for Scenario 1, the computed confidence intervals contain zero; hence hypothesis 2.3 is accepted. In other words, we conclude that the probability vectors of the two populations are equal, consistent with the simulation settings.

In Scenario 2, except for the confidence intervals of $\beta^{(1)}$ and $D^{(1)}$, all other confidence intervals contain zero, leading us to correctly reject hypothesis 2.3. Specifically, we can observe that the probabilities for the first group of the two trinomial populations are not equal, while the probabilities for the second group are, which is consistent with Scenario 2.

In Scenario 3, all confidence intervals except for $\mu^{(2)}$ do not contain zero, leading us to conclude that the probabilities for both groups of the two trinomial populations are not equal.

# 6. A Real-World Example: Population Aged 15 and Older by Economic Activity Status from 2009 to 2021

In this section, we study the proposed model on a real dataset, examining the population aged 15 and older in the Islamic Republic of Iran between 2009 and

Table 2: *Estimates, Mean Squared Errors, and Confidence Intervals for* $D^{(1)}$ *and* $D^{(2)}$

| | | $D^{(1)}$ | $D^{(2)}$ |
|---|---|---|---|
| Estimate | Scenario 1 | 0.0032 | $-0.0313$ |
| | Scenario 2 | 0.1542 | $-1.7120$ |
| | Scenario 3 | 0.4856 | $-0.3575$ |
| MSE | Scenario 1 | $1.08 \times 10^{-5}$ | 0.0009 |
| | Scenario 2 | 0.0238 | 2.9327 |
| | Scenario 3 | 0.2360 | 0.1288 |
| Confidence Interval | Scenario 1 | $(-0.0475, 0.0577)$ | $(-0.0911, 0.0068)$ |
| | Scenario 2 | $(0.0783, 0.2319)$ | $(-6.9223, 0.0532)$ |
| | Scenario 3 | $(0.3409, 0.5851)$ | $(-0.6307, -0.0146)$ |

Table 3: *Estimates, Mean Squared Errors, and Confidence Intervals for* $\mu^{(1)}$ *and* $\mu^{(2)}$

| | | $\mu^{(1)}$ | $\mu^{(2)}$ |
|---|---|---|---|
| Estimate | Scenario 1 | $-0.0595$ | 0.0183 |
| | Scenario 2 | 0.0132 | $-1.0729$ |
| | Scenario 3 | 0.7236 | $-0.6491$ |
| MSE | Scenario 1 | 0.0078 | 0.0033 |
| | Scenario 2 | 0.0046 | 1.1555 |
| | Scenario 3 | 0.5240 | 0.4229 |
| Confidence Interval | Scenario 1 | $(-1.1321, 1.3267)$ | $(-1.3401, 1.3686)$ |
| | Scenario 2 | $(-1.7783, 1.4935)$ | $(-4.9956, 1.1256)$ |
| | Scenario 3 | $(0.1537, 2.4793)$ | $(-3.6337, 0.1489)$ |

2021. The data is available in Table 4 and can be accessed from the website of the Statistical Center of Iran. To ensure uniformity in the data across the two periods, adjustments were made to account for changes in provincial boundaries by merging the provinces of Tehran and Alborz for the year 2021.

The population aged 15 and older is categorized into employed individuals, unemployed individuals, and economically inactive individuals. Therefore, the dataset represents two trinomial populations: the population aged 15 and older in 2021 as the first population, and the active population in 2009 as the second population. Consequently, $D_1$ represents the logarithm of the odds ratio of employment in 2021 to 2009, while $D_2$ represents the logarithm of the odds ratio of economically inactive individuals in 2021 to that group in 2009.

The analysis is conducted across 30 provinces, considered as study centers. Additionally, the ratio of literate to illiterate individuals in 2021 is included as an auxiliary variable for $D_1$ and $D_2$, respectively. Following the simulation protocol, the Gibbs sampler is executed 1500 times, with the first 500 executions regarded as the burn-in period.

The estimate for $D_1$ is 0.046, with a 95% confidence interval of $(-0.055, 0.011)$. Since the confidence interval encompasses zero, we fail to reject the null hypothesis $H_0: D_1 = 0$. This suggests that the ratio of employed individuals in 2009 and 2021 is statistically equivalent. Similarly, $\beta_1$, the coefficient for the auxiliary variable, is 0.0005, with a 95% confidence interval of $(-0.0006, 0.0013)$, which also includes zero. This indicates no significant association between the employment odds ratio and the ratio of literate individuals.

These findings imply that over the 12-year period, there was no measurable change in the relative proportion of employment among the population aged 15 and older. Moreover, the literacy ratio did not emerge as a significant factor influencing this trend. This result could suggest stability in the employment structure across provinces, potentially reflecting a lack of transformative economic or policy interventions during this time frame.

The estimate for $D_2$ is $-0.953$, with a 95% confidence interval of $(-2.895, 0.213)$, which also includes zero. Thus, we fail to reject the null hypothesis $H_0: D_2 = 0$, suggesting no significant difference in the ratio of economically inactive individuals between 2009 and 2021. The auxiliary variable for $D_2$, the ratio of illiterate individuals, also showed no significant association, with $\beta_2 = -0.071$ and a 95% confidence interval of $(-0.218, 0.186)$.

The lack of significant change in $D_2$ suggests that the proportion of economically inactive individuals has remained stable over time. This result, combined with the findings for $D_1$, could indicate broader socio-economic trends in Iran, such as persistent barriers to workforce participation or the inability of economic

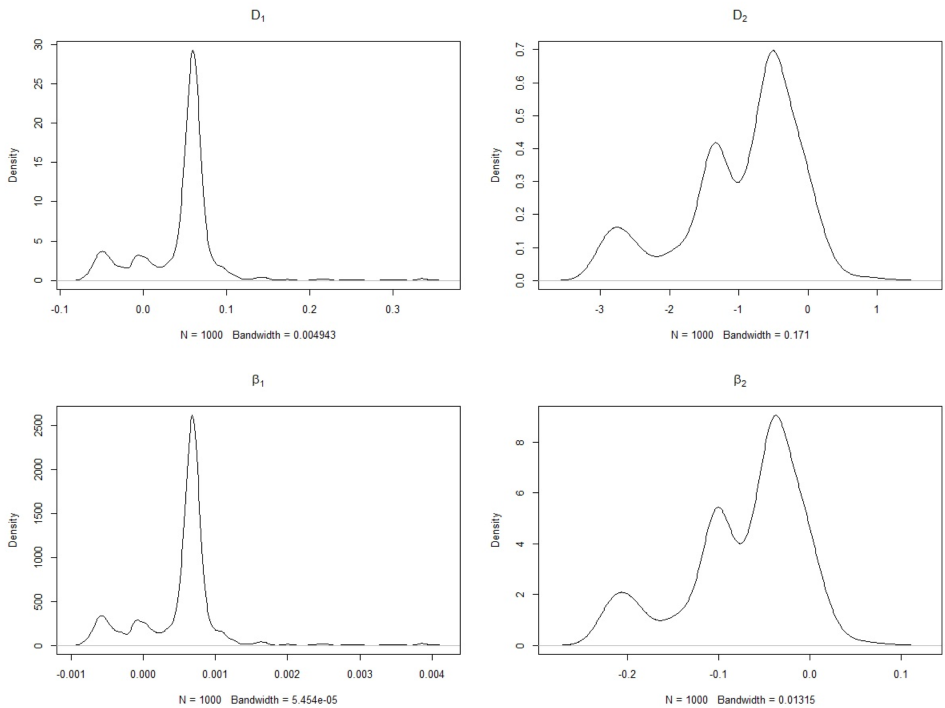


Figure 1: Densities of $D_1$, $D_2$, $\beta_1$, and $\beta_2$

growth to translate into improved employment opportunities. The insignificance of the literacy ratio as a predictor highlights the potential need for more targeted policies addressing the underlying factors of economic inactivity, such as skill mismatches or regional disparities.

The densities of $D_1$, $D_2$, $\beta_1$, and $\beta_2$ are plotted in Figure 1. These results underscore the importance of employing nuanced statistical models like the one proposed in this paper to gain deeper insights into socio-economic trends and their implications for policy-making.

# 7. Results and Discussion

In this study, we employed Bayesian meta-regression models to analyze treatment effects across multiple clinical studies, effectively addressing heterogeneity and confounding variables. Utilizing a semiparametric Bayesian framework, we incorporated auxiliary variables to account for variability among studies. We conducted simulations across three datasets, each reflecting different levels of heterogeneity in treatment response, and estimated posterior distributions of treatment effects using Gibbs sampling and the Metropolis-Hastings algorithm.

Our model was also applied to a real dataset, illustrating its practical applicability. The results from both simulations and real-world applications highlight the advantages of Bayesian meta-regression in clinical research. By incorporating prior information and managing heterogeneity, these models yield more accurate and interpretable results than traditional methods. Our findings align with the assertions of Thompson (1994) and Burr and Doss (2005) that understanding heterogeneity enhances the interpretability of treatment effects and supports evidence-based decision-making in healthcare.

The flexibility of our framework in handling complex data structures positions it as a valuable tool for future meta-analytic studies. As clinical research evolves, the integration of Bayesian methods will become increasingly crucial for synthesizing evidence across diverse studies and informing clinical practice. In conclusion, our study emphasizes the need for advanced statistical techniques, such as Bayesian meta-regression, to tackle the challenges of heterogeneity in clinical trials. Future research should consider extending these models to contexts like longitudinal studies and multi-arm trials to enhance their applicability and impact on healthcare decision-making.

## Acknowledgement

We would like to express our sincere gratitude to the reviewers for their thorough and insightful feedback on this manuscript. We appreciate the time and effort they invested in providing constructive critiques and enhancing the clarity and rigor of this research.

# Proof of theorem 3.1

By combining Equations 3.11 and 3.13, we can derive the posterior distribution of $\eta_i^{(1)}, \eta_i^{(2)}$ conditional on $\eta_{(-i)}^{(1)}, \eta_{(-i)}^{(2)}, \boldsymbol{\alpha}$ as follows:

$$\begin{aligned}
&\pi_D(\eta_i^{(1)}, \eta_i^{(2)} | \eta_{(-i)}^{(1)}, \eta_{(-i)}^{(2)}, \boldsymbol{\alpha}) \propto \\
&\qquad \pi(\eta_i^{(1)}, \eta_i^{(2)} | \eta_{(-i)}^{(1)}, \eta_{(-i)}^{(2)}, \boldsymbol{\alpha}) L(D | \eta_i^{(1)}, \eta_i^{(2)}, \eta_{(-i)}^{(1)}, \eta_{(-i)}^{(2)}, \boldsymbol{\alpha}) \\
&\qquad \propto \pi(\eta_i^{(1)} | \eta_{(-i)}^{(1)}, \mu^{(1)}, \tau^{(1)}) \pi(\eta_i^{(2)} | \eta_{(-i)}^{(2)}, \mu^{(2)}, \tau^{(2)}) \\
&\qquad\quad \times L(D | \eta_i^{(1)}, \eta_i^{(2)}, \eta_{(-i)}^{(1)}, \eta_{(-i)}^{(2)}, \boldsymbol{\alpha}) \\
&\qquad \propto \left[ \frac{M N_-^{\mu^{(1)}}(\mu^{(1)}, \tau^{2(1)}) + \sum_{k \neq i, \eta_k^{(1)} < \mu^{(1)}} \delta_{\eta_k^{(1)}}}{\frac{M}{2} + m_-^{(1)}} + \frac{M N_+^{\mu^{(1)}}(\mu^{(1)}, \tau^{2(1)}) + \sum_{k \neq i, \eta_k^{(1)} > \mu^{(1)}} \delta_{\eta_k^{(1)}}}{\frac{M}{2} + m_+^{(1)}} \right] \\
&\qquad\quad \times \left[ \frac{M N_-^{\mu^{(2)}}(\mu^{(2)}, \tau^{2(2)}) + \sum_{k \neq i, \eta_k^{(2)} < \mu^{(2)}} \delta_{\eta_k^{(2)}}}{\frac{M}{2} + m_-^{(2)}} + \frac{M N_+^{\mu^{(2)}}(\mu^{(2)}, \tau^{2(2)}) + \sum_{k \neq i, \eta_k^{(2)} > \mu^{(2)}} \delta_{\eta_k^{(2)}}}{\frac{M}{2} + m_+^{(2)}} \right] \\
&\qquad\qquad\qquad \times N_2(x_i'^{(1)} \eta_i^{(1)}, x_i'^{(2)} \eta_i^{(2)}, \sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i)
\end{aligned}$$

$$\begin{aligned}
\pi_D(\eta_i^{(1)}, \eta_i^{(2)} | \eta_{(-i)}^{(1)}, \eta_{(-i)}^{(2)}, \boldsymbol{\alpha}) \propto & \left[ \frac{M^2 N_-^{\mu^{(1)}}(\mu^{(1)}, \tau^{2(1)}) N_-^{\mu^{(2)}}(\mu^{(2)}, \tau^{2(2)})}{\frac{M^2}{4} + \frac{M}{2}(m_-^{(1)} + m_-^{(2)}) + m_-^{(1)} m_-^{(2)}} \right] \\
& \qquad N_2(x_i'^{(1)} \eta_i^{(1)}, x_i'^{(2)} \eta_i^{(2)}, \sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i) \\
& + \left[ \frac{M^2 N_-^{\mu^{(1)}}(\mu^{(1)}, \tau^{2(1)}) N_+^{\mu^{(2)}}(\mu^{(2)}, \tau^{2(2)})}{\frac{M^2}{4} + \frac{M}{2}(m_-^{(1)} + m_+^{(2)}) + m_-^{(1)} m_+^{(2)}} \right] N_2(x_i'^{(1)} \eta_i^{(1)}, x_i'^{(2)} \eta_i^{(2)}, \sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i) \\
& + \left[ \frac{M^2 N_+^{\mu^{(1)}}(\mu^{(1)}, \tau^{2(1)}) N_-^{\mu^{(2)}}(\mu^{(2)}, \tau^{2(2)})}{\frac{M^2}{4} + \frac{M}{2}(m_+^{(1)} + m_-^{(2)}) + m_+^{(1)} m_-^{(2)}} \right] N_2(x_i'^{(1)} \eta_i^{(1)}, x_i'^{(2)} \eta_i^{(2)}, \sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i) \\
& + \left[ \frac{M^2 N_+^{\mu^{(1)}}(\mu^{(1)}, \tau^{2(1)}) N_+^{\mu^{(2)}}(\mu^{(2)}, \tau^{2(2)})}{\frac{M^2}{4} + \frac{M}{2}(m_+^{(1)} + m_+^{(2)}) + m_+^{(1)} m_+^{(2)}} \right] N_2(x_i'^{(1)} \eta_i^{(1)}, x_i'^{(2)} \eta_i^{(2)}, \sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i) \\
& + \zeta \\
= & \; I + II + III + IV + \zeta.
\end{aligned}$$

Here, $\zeta$ represents the remaining terms of the posterior distribution introduced in the theorem, which include the point masses. We now calculate each term of the above equation separately.

$$
\begin{aligned}
I &= \left[\frac{M^2 N_-^{\mu^{(1)}}(\mu^{(1)}, \tau^{2(1)}) N_-^{\mu^{(2)}}(\mu^{(2)}, \tau^{2(2)})}{\frac{M^2}{4} + \frac{M}{2}(m_-^{(1)} + m_-^{(2)}) + m_-^{(1)} m_-^{(2)}}\right] N_2(x_i'^{(1)}\eta_i^{(1)}, x_i'^{(2)}\eta_i^{(2)}, \sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i) \\
&= \left[\frac{M^2 I_{(-\infty,\mu^{(1)})}(\eta_i^{(1)}) I_{(-\infty,\mu^{(2)})}(\eta_i^{(2)})}{\frac{M^2}{4} + \frac{M}{2}(m_-^{(1)} + m_-^{(2)}) + m_-^{(1)} m_-^{(2)}}\right] N^{\mu^{(1)}}(\mu^{(1)}, \tau^{2(1)}) N^{\mu^{(2)}}(\mu^{(2)}, \tau^{2(2)}) \\
&\qquad \times N_2(x_i'^{(1)}\eta_i^{(1)}, x_i'^{(2)}\eta_i^{(2)}, \sigma_i^{2(1)}, \sigma_i^{2(2)}, \rho_i) \\
&= \left[\frac{M^2 I_{(-\infty,\mu^{(1)})}(\eta_i^{(1)}) I_{(-\infty,\mu^{(2)})}(\eta_i^{(2)})}{\frac{M^2}{4} + \frac{M}{2}(m_-^{(1)} + m_-^{(2)}) + m_-^{(1)} m_-^{(2)}}\right] \\
&\quad \times \frac{1}{4\pi^2\sqrt{(1-\rho_i^2)}\sigma_i^{(1)}\sigma_i^{(2)}\tau^{(1)}\tau^{(2)}} \exp\left[\frac{-\left(\eta_i^{(1)} - \mu^{(1)}\right)^2}{2\tau^{2(1)}}\right] \\
&\quad \times \exp\left[-\frac{1}{2(1-\rho_i^2)}\left\{\left(\frac{x_i'^{(1)}\beta_i^{(1)} - \eta_i^{(1)} x_i'^{(1)}}{\sigma_i^{(1)}}\right)^2 + \left(\frac{x_i'^{(2)}\beta_i^{(2)} - \eta_i^{(2)} x_i'^{(2)}}{\sigma_i^{(2)}}\right)^2\right\}\right] \\
&\quad \times \exp\left[\frac{\rho_i}{1-\rho_i^2}\left(\frac{x_i'^{(1)}\beta_i^{(1)} - \eta_i^{(1)} x_i'^{(1)}}{\sigma_i^{(1)}}\right)\left(\frac{x_i'^{(2)}\beta_i^{(2)} - x_i'^{(2)}\eta_i^{(2)}}{\sigma_i^{(2)}}\right) - \frac{\left(\eta_i^{(2)} - \mu^{(2)}\right)^2}{2\tau^{2(2)}}\right]
\end{aligned}
$$

Using some straightforward algebraic calculations, we have:

$$
\begin{aligned}
I &= \left[\frac{M^2 I_{(-\infty,\mu^{(1)})}(\eta_i^{(1)}) I_{(-\infty,\mu^{(2)})}(\eta_i^{(2)})}{\frac{M^2}{4} + \frac{M}{2}(m_-^{(1)} + m_-^{(2)}) + m_-^{(1)} m_-^{(2)}}\right] \frac{1}{4\pi^2\sqrt{(1-\rho_i^2)}\sigma_i^{(1)}\sigma_i^{(2)}\tau^{(1)}\tau^{(2)}} \\
&\quad \times \exp\left(-\frac{1}{2(1-\rho_i^2)}\left\{\left(\frac{\eta_i^{(1)} - A_i^{(1)}}{B_i^{(1)}}\right)^2 + \left(\frac{\eta_i^{(2)} - A_i^{(2)}}{B_i^{(2)}}\right)^2\right\}\right) \\
&\quad \times \exp\left(R - \frac{1}{2(1-\rho_i^2)}\left\{-2\rho_i \frac{1}{B_i^{(1)} B_i^{(2)}}(\eta_i^{(1)} - A_i^{(1)})(\eta_i^{(2)} - A_i^{(2)})\right\}\right) \\
&\quad \times \exp\left[-\frac{\left(x_i'^{(1)}\beta_i^{(1)} - x_i'^{(1)}\mu^{(1)}\right)}{2\left[(1-\rho_i^2)\sigma_i^{2(1)} + x_i'^{(1)}\tau^{2(1)} x_i^{(1)}\right]} - \frac{\left(x_i'^{(2)}\beta_i^{(2)} - x_i'^{(2)}\mu^{(2)}\right)}{2\left[(1-\rho_i^2)\sigma_i^{2(2)} + x_i'^{(2)}\tau^{2(2)} x_i^{(2)}\right]}\right] \\
&= C_{-,-}^{(1,2)} N_{-,-}^{(\mu^{(1)},\mu^{(2)})}(A_i^{(1)}, A_i^{(2)}, B_i^{2(1)}, B_i^{2(2)}, \rho_i)
\end{aligned}
$$

Similarly, it can be shown that:

$$
\begin{aligned}
II &= C_{-,+}^{(1,2)} N_{-,+}^{(\mu^{(1)},\mu^{(2)})}(A_i^{(1)}, A_i^{(2)}, B_i^{2(1)}, B_i^{2(2)}, \rho_i) \\
III &= C_{+,-}^{(1,2)} N_{+,-}^{(\mu^{(1)},\mu^{(2)})}(A_i^{(1)}, A_i^{(2)}, B_i^{2(1)}, B_i^{2(2)}, \rho_i) \\
IV &= C_{+,+}^{(1,2)} N_{+,+}^{(\mu^{(1)},\mu^{(2)})}(A_i^{(1)}, A_i^{(2)}, B_i^{2(1)}, B_i^{2(2)}, \rho_i)
\end{aligned}
$$

## References

Antoniak, C.E. (1974). Mixtures of Dirichlet Processes with Applications to Bayesian Nonparametric Problems. *The Annals of Statistics*, **2**, 1152-1174.

Armitage, P. and Colton, T. (1998). Encyclopedia of Biostatistics. *John Wiley and Sons*, 1st ed., Vol. 4, Inc.: New York.

Bashore, T.R., Osman. A. and Heffley, E.F. (1989). Mental slowing in elderly persons: a cognitive psychophysiological analysis. *Psychology and Aging*, **4**, 235-244.

Burr, D. and Doss, H. (2005). A Bayesian Semi-parametric Model for Random Effects Meta-Analysis. *Journal of the American Statistical Association*, **100**, 242-251.

Carlin, J. B. (1992). Meta-Analysis for 2 $\times$ 2 Tables: A Bayesian Approach. *Journal of Statistics in Medicine*, **11**, 141-158.

Carota, C. and Parmigiani, G. (2002). Semiparametric regression for count data. *Biometrika*, **89**, 265-281.

Chung, Y. and Dunson, D.B. (2007). Local Dirichlet Process. *Working paper, Duke University.*

Dominici, F. and Parmigiani, G. (2001). Bayesian Semiparametric Analysis of Developmental Data. *Biometrics*, **57**, 150-157.

Doss, H. (1985). Bayesian Nonparametric Estimation of the Median: Part I: Computation of the Estimates. *The Annals of Statistics*, **13**, 1432-1444.

Ferguson, T. S. (1973). A Bayesian Analysis of Some Nonparametric Problems. *The Annals of Statistics*, **1**, 209-230.

Frömke, C., Kirstein, M., and Zapf, A. (2022). A semiparametric approach for meta-analysis of diagnostic accuracy studies with multiple cut-offs. Research Synthesis Methods, **13(5)**, 612-621.

Greenland, S. (1994). A critical look at some popular meta-analytic methods. *American Journal of Epidemiology*, **140**, 290-296.

Griffin, J.E., and Steel, M.F.J., (2007). Order-Based Dependent Dirichlet Processes. *Working paper, Duke University.*

Higgins, J.P.T. (1997). Exploiting Information in Random Effects Meta-Analysis. *Ph.D. Thesis, University of Reading.*

Jones, D.R. (1992). Meta-analysis of observational epidemiological studies: a review. *Journal of the Royal Society of Medicine*, **85**, 165-168.

Maier, M., Bartoš, F., and Wagenmakers, E.-J. (2022). Robust Bayesian meta-analysis: Addressing publication bias with model-averaging. Psychological Methods. Advance online publication.

Ohlsen, D.I, Sharples, L.D. and Spiegelhalter, D.J. (2007). Flexible random effect models using Bayesian semi-parametric models: applications to institutional comparisons. *Journal of Statistics in Medicine*, **26**, 2088-2112.

Pearson, K. (1933). On A Method of Determining Whether A Sample Size n Supposed To Have Been Drawn from a Parent Population Having A Known Probability Integral Has Property Been Drawn At Random. *Biometrika*, **25**, 379-410.

Thompson, S.G. (1994). Why sources of heterogeneity in meta-analysis should be investigated. *British Medical Journal*, **309**, 1351-1355.

Tippett, L.H.C. (1931). The Methods of Statistics. *London: Williams and Norgate.*

Table 4: Population aged 15 and older by economic activity status by province

| Year 1400 | | | | Year 1388 | | | |
|---|---|---|---|---|---|---|---|
| Literacy Rate (%) | Economically Inactive | Economically Active | | Economically Inactive | Economically Active | | Province |
| | | Employed | Unemployed | | Unemployed | Employed | |
| 85.7 | 1802192 | 1187975 | 97795 | 1491012 | 145662 | 1289921 | East Azerbaijan |
| 81.9 | 1428491 | 974062 | 137676 | 1111680 | 116548 | 960584 | West Azerbaijan |
| 83.3 | 525669 | 405526 | 46479 | 485892 | 61100 | 443253 | Ardabil |
| 90.0 | 2451538 | 1522156 | 180009 | 2018530 | 201326 | 1474925 | Isfahan |
| 84.1 | 295737 | 141814 | 11430 | 226338 | 22426 | 154579 | Ilam |
| 91.1 | 539095 | 305535 | 30186 | 394812 | 31855 | 239443 | Bushehr |
| 92.8 | 7945589 | 5071645 | 421232 | 6582574 | 552981 | 4090467 | Tehran and Alborz |
| 85.5 | 417457 | 261146 | 36666 | 369207 | 47346 | 241006 | Chaharmahal and Bakhtiari |
| 87.3 | 323420 | 217260 | 17035 | 240473 | 14966 | 196448 | South Khorasan |
| 90.1 | 2928720 | 1895519 | 133644 | 2405643 | 233002 | 1813961 | Razavi Khorasan |
| 82.3 | 345013 | 251329 | 27150 | 331605 | 15240 | 221983 | North Khorasan |
| 87.9 | 2126706 | 1236139 | 178313 | 1846586 | 176489 | 1172274 | Khuzestan |
| 86.2 | 447276 | 348892 | 24413 | 384481 | 30927 | 333166 | Zanjan |
| 91.6 | 360191 | 197524 | 15201 | 274320 | 16051 | 160678 | Semnan |
| 77.5 | 1215199 | 580643 | 60278 | 907551 | 70055 | 452077 | Sistan and Baluchestan |
| 90.3 | 2400406 | 1318166 | 109006 | 2022105 | 198651 | 1169440 | Fars |
| 89.8 | 566401 | 419240 | 32326 | 557944 | 44139 | 316422 | Qazvin |
| 89.8 | 639440 | 340278 | 34691 | 504995 | 33587 | 278102 | Qom |
| 81.4 | 723510 | 467159 | 62096 | 628947 | 51907 | 428025 | Kurdistan |
| 89.2 | 1438779 | 862304 | 108947 | 1055433 | 78894 | 586572 | Kerman |
| 85.0 | 874145 | 561130 | 98285 | 964140 | 87651 | 544068 | Kermanshah |
| 84.8 | 348484 | 167136 | 16813 | 302386 | 22288 | 129137 | Kohgiluyeh and Boyer-Ahmad |
| 85.4 | 884676 | 500819 | 41045 | 700280 | 39868 | 519926 | Golestan |
| 88.1 | 1165466 | 871254 | 88746 | 1117182 | 137890 | 757894 | Gilan |
| 83.6 | 782438 | 481520 | 68658 | 796499 | 109773 | 464615 | Lorestan |
| 89.9 | 1554374 | 1098178 | 82978 | 1402282 | 75730 | 859214 | Mazandaran |
| 87.7 | 722874 | 384963 | 32660 | 602616 | 50093 | 394766 | Markazi |
| 88.3 | 751349 | 488177 | 91849 | 612694 | 24643 | 329970 | Hormozgan |
| 85.3 | 779551 | 519286 | 39679 | 749325 | 110131 | 499098 | Hamedan |
| 92.5 | 459219 | 370677 | 49669 | 442730 | 30598 | 307002 | Yazd |