Communications for Statistical Applications and Methods
2022, Vol. 29, No. 5, 547–559

https://doi.org/10.29220/CSAM.2022.29.5.547
Print ISSN 2287-7843 / Online ISSN 2383-4757

A pooled Bayes test of independence using restricted pooling model for contingency tables from small areas

Aejeong Jo[a] , Dal Ho Kim[1,b]

[a]Division of Research & Development, National Evidence-based Healthcare Collaborating Agency, Korea; [b]Department of Statistics, Kyungpook National University, Korea

Abstract

For a chi-squared test, which is a statistical method used to test the independence of a contingency table of two factors, the expected frequency of each cell must be greater than 5. The percentage of cells with an expected frequency below 5 must be less than 20% of all cells. However, there are many cases in which the regional expected frequency is below 5 in general small area studies. Even in large-scale surveys, it is difficult to forecast the expected frequency to be greater than 5 when there is small area estimation with subgroup analysis. Another statistical method to test independence is to use the Bayes factor, but since there is a high ratio of data dependency due to the nature of the Bayesian approach, the low expected frequency tends to decrease the precision of the test results. To overcome these limitations, we will borrow information from areas with similar characteristics and pool the data statistically to propose a pooled Bayes test of independence in target areas. Jo *et al.* (2021) suggested hierarchical Bayesian pooling models for small area estimation of categorical data, and we will introduce the pooled Bayes factors calculated by expanding their restricted pooling model. We applied the pooled Bayes factors using bone mineral density and body mass index data from the Third National Health and Nutrition Examination Survey conducted in the United States and compared them with chi-squared tests often used in tests of independence.

Keywords: pooled Bayes factor, nonparametric hierarchical Bayesian model, Dirichlet process, slice sampling, small area

1. Introduction

When the contingency table consists of two multinomial variables, we test independence and obtain measures of the association among variables. Particularly, researchers in clinical studies are highly interested in the test because the independence test result reflects the simple relationship between two different clinical factors. As a simple example, we can consider the hypothesis that smoking can cause lung cancer. Smoking status is generally divided into three categories: never smoking, previous smoking, or current smoking. Before the main analysis for smoking status and incidence of lung cancer, a study performed a test of independence. Although the test does not show the direction between the variables, it is important to show the possibility of the study. If an association is found between two variables, there will be a significant difference between the occurrence rates of lung cancer in the three categories of smoking. Naturally, clinical interest is crucial at both county and state levels. However, such studies have a drawback in terms of the power of the hypothesis test

[1] Corresponding author: Department of Statistics, Kyungpook National University, 80 Daehakro, Bukgu, Deagu 41566, Korea. E-mail: dalkim@knu.ac.kr

Published 30 September 2022 / journal homepage: http://csam.or.kr

because the sample size is much smaller than the unit of the state. In other words, as the sample size decreases, power decreases.

Over the years, various attempts have been made to solve this problem. One of these solutions is pooling strategies for data from small areas. The key to the pooling method is estimating the test statistic used to borrow strength from neighboring areas to increase the stat power. Consonni *et al.* (1995), DuMouchel *et al.* (1983). Evans and Sedransk (2003) developed the fully Bayesian justification using the geometric means of Malec and Sedransk (1992), who proposed a Bayesian model to estimate the parameter mean corresponding to the same specified experiment between similar experiments. Subsequently, studies have applied the Dirichlet process proposed by Ferguson (1973) to pool the data. Donson (2009) developed a generalization of the Dirichlet process prior, in which the model allows dependent local pooling and borrowing of information from data with similar experiments. Specifically, the Dirichlet process prior suggests using the slice sampling proposed by Walker (2007). Nandram *et al.* (2019) developed a pooled Bayes test of independence for sparse contingency tables based on the idea of generalization developed by Donson (2009). They constructed the hierarchical Bayesian model with Dirichlet distribution before the test of independence between Bone Mineral Density (BMD) and Body Mass Index (BMI) from the Third National Health and Nutrition Examination Survey (NHANES III) in the U.S. In their model, the pooling of information was implemented from the hyperprior of an interesting parameter.

Another issue of the pooled Bayes factor (PBF) is the integration of the joint density function to calculate the marginal likelihood. The joint density function is very complex and impossible to integrate. Therefore, the equation is computed by Monte Carlo integration using the new method proposed by Nandram and Kim (2002). Their method utilizes the multiplication rule of probability to use a hierarchical Bayesian structure. The key of their new method is that the importance function is considered with the conditional posterior density of each variable. Nandram *et al.* (2019) further suggested a pooled Bayes test of independence for contingency tables. They developed a Bayesian model using the Dirichlet multinomial hierarchical Bayesian model proposed by Nandram (1998). Recently, Jo *et al.* (2021) developed a pooled hierarchical Bayesian model to analyze categorical data from small areas using the Dirichlet multinomial distribution. They proposed the hierarchical Bayesian model that borrows information from similar regions to pool the data, estimated the finite population proportion, and compared it for each model.

This study constructs hierarchical Bayesian models for the pooled Bayes factor based on the pooled test of independence proposed by Nandram *et al.* (2019). We expanded the restricted pooling model proposed by Jo *et al.* (2021) to calculate the pooled Bayes factor and compared it with the no pooling model, which has no pooling of data, and the complete pooling model that considers all the regions as one group. First, we construct simple Bayesian models with two extreme cases: a completely pooled Bayesian model versus a no-pooled model. We can then identify the characteristics of pooling from two basic models. Additionally, basic models may be expanded in the Bayesian pooling model. We call this model the restricted-pooling model and its nonparametric version. To compare the performance of the pooled Bayes factor, we analyze the association between BMD and BMI. The BMI and BMD data are from the MHANES III survey data.

The association between the BMI and BMD variables is a steadily important issue in clinical studies. The BMI is a more easily comprehensible measure of obesity and is divided into four levels: (1) Underweight: BMI less than 20 kg/m^2; (2) Optimal: BMI from 20 kg/m^2 to 25 kg/m^2; (3) Overweight: BMI from 25 kg/m^2 to 30 kg/m^2, and (4) Obese: BMI above 30 kg/m^2. Additionally, BMD is an important indicator of clinical prognoses, such as osteopenia, osteoporosis, and fracture. It is measured according to optical density per cm^2 of the bone surface upon medical imaging and is categorized into

three levels: (1) normal: BMD less than 1 standard deviation (SD) below the young non-Hispanic white (NHW) adult mean, (2) osteopenia: BMD from 1 to 2.5 SD below the young NHW adult mean, (3) osteoporosis - BMD more than 2.5 SD below the young NHW adult mean. This study analyzes the BMI and BMD data using various pooled Bayes factors with hierarchical Bayesian models and compares the performance and results of these models. In Section 2, we introduce hierarchical Bayesian models for the pooled Bayes factors. In Section 3, we propose a nonparametric version of several models. In Section 4, we present the results of the data analysis concerning the BMI and BMD data. Finally, in Section 5, we discuss the eight models.

2. Hierarchical Bayesian parameteric models

For the s^{th} area of S small areas, we consider the $r \times c$ contingency tables with cell counts, n_{jk}, which are the responses for the k^{th} column and j^{th} row in the area. Let π_{sjk} denote the corresponding probabilities of each unit cell in the s^{th} area. When p_{sj} and q_{sk} are marginal probabilities for each column and row in the s^{th} area, the independent models have $\pi_{sjk} = p_{sj}q_{sk}$, $j = 1, \ldots, r$, $k = 1, \ldots, c$, $\sum_{j=1}^{r} p_{sj} = 1$, and $\sum_{k=1}^{c} p_{sk} = 1$ for $s = 1, \ldots, S$.

2.1. General models

Let n_{si}, $i = 1, \ldots, I(= rc)$ be the cell counts for the sth area and π_{si} denote the corresponding probabilities of each area. We assume that

$$\mathbf{n}_s|\boldsymbol{\pi}_s \overset{\text{ind}}{\sim} \text{Multi}(n_{s\cdot}, \boldsymbol{\pi}_s), \quad s = 1, \ldots, S, \tag{2.1}$$

where $\mathbf{n}_s = (n_{s1}, \ldots, n_{sI})$ for $s = 1, \ldots, S$ is the vector of responses, $n_{s\cdot} = \sum_{i=1}^{I} n_{si}$ is total sum of responses, $\boldsymbol{\pi}_s = (\pi_{s1}, \ldots, \pi_{sI})$ is the corresponding probability vector of each area, and $\sum_{i=1}^{I} \pi_{si} = 1$. In here, I is defined the number of cells for the table corresponding to each area. We construct the hierarchical Bayesian model with a mixture of Dirichlet prior, which comprises two parts. The effective part of the modeling is distinguished from the uncertainty part. We are interested in estimating the parameter from the data, excluding the uncertainty of the model. This is called the restricted pooling model and is as follows.

$$\begin{aligned}
\boldsymbol{\pi}_s|\boldsymbol{\mu}, \tau &\sim \phi\text{Dirichlet}(\boldsymbol{\mu}\tau) + (1-\phi)\text{Dirichlet}(\mathbf{1}),\\
\pi(\boldsymbol{\mu}, \tau) &= \frac{(I-1)!}{(1+\tau)^2}, \quad \tau > 0,\\
\phi &\sim \text{Uniform}\left(\frac{1}{2}, 1\right),
\end{aligned}$$

where $\mathbf{1}$ is $I \times 1$ dimensional vector with all values equal to 1, ϕ is from a uniform distribution with domain $(1/2, 1)$ and determines the proportion of uncertainty in the entire data. We assume that the uncertain region is smaller than the effectiveness through a prior distribution of ϕ. The parameters $\boldsymbol{\mu}$,τ, and ϕ are mutually independent, and the posterior joint density is

$$\pi(\boldsymbol{\pi}, \boldsymbol{\mu}, \tau, \phi|\mathbf{n}) = \prod_{s=1}^{S}\left[\frac{n_s!}{\prod_{i=1}^{I} n_{si}!}\prod_{i=1}^{I}\pi_{si}^{n_{si}}\left\{\phi\frac{1}{D(\boldsymbol{\mu}\tau)}\prod_{i=1}^{I}\pi_{si}^{\mu_i\tau-1} + (1-\phi)(I-1)\right\}\right]\frac{(I-1)!}{(1+\tau)^2},$$

where $\mathbf{n} = (\mathbf{n}_1, \ldots, \mathbf{n}_S)$, $\boldsymbol{\pi} = (\boldsymbol{\pi}_1, \ldots, \boldsymbol{\pi}_S)$, $D(\boldsymbol{\mu}\tau) = \prod_{i=1}^{I} \Gamma(\mu_i\tau)/\Gamma(\sum_{i=1}^{I} \mu_i\tau)$ is the multivariate Beta function. Then the marginal likelihood in the sth area is

$$f(\mathbf{n}_s) = \frac{(I-1)!n_s!}{\prod_{i=1}^{I} n_{si}!} \iint \frac{1}{(1+\tau)^2} \left\{ \frac{3}{4} \frac{D(\boldsymbol{\mu}\tau + \mathbf{n}_s)}{D(\boldsymbol{\mu}\tau)} + \frac{1}{4}(I-1)!D(\mathbf{n}_s + 1) \right\} d\tau d\boldsymbol{\mu}.$$

We will now calculate the marginal likelihood using the properties of the definite integral of a linear function for ϕ and the posterior for $\boldsymbol{\pi}$. The posterior density of $\boldsymbol{\pi}_s$ for $s = 1, \ldots, S$ is

$$\boldsymbol{\pi}_s \mid \mathbf{n}, \boldsymbol{\mu}, \tau \sim \text{Dirichlet}(\mathbf{n}_s + \boldsymbol{\mu}\tau)$$

for the formar of integral equation and

$$\boldsymbol{\pi}_s \mid \mathbf{n}, \boldsymbol{\mu}, \tau \sim \text{Dirichlet}(\mathbf{n}_s + \mathbf{1}).$$

Additionally, we can apply the new method developed by Nandram and Kim (2002).

2.2. Independence models

Let n_{sjk}, $j = 1, \ldots, r, k = 1, \ldots, c$, be the cell counts for j^{th} row and kth column in s^{th} area, $s = 1, \ldots, S$ with corresponding cell probability $\pi_{sjk} = p_{sj}q_{sk}$ where $p_{sj} = \sum_{k=1}^{c} \pi_{sjk}$ and $q_{sk} = \sum_{j=1}^{r} \pi_{sjk}$. For our pooled Bayes factor of independence, we assume that

$$\mathbf{n}_s | \boldsymbol{p}_s, \boldsymbol{q}_s \overset{\text{ind}}{\sim} \text{Multi}(n_s, \text{vec}(\boldsymbol{p}_s\boldsymbol{q}_s')), \quad s = 1, \ldots, S,$$

where $\mathbf{n}_s = (n_{s11}, \ldots, n_{s1c}, \ldots, n_{sr1}, \ldots, n_{src})$, $n_s = \sum_{j=1}^{r}\sum_{k=1}^{c} n_{sjk}$, $\boldsymbol{p}_s = (p_{s1}, \ldots, p_{sr})$, $\boldsymbol{q} = (q_{s1}, \ldots, q_{sc})$, $\sum_{j-1}^{r} p_{sj} = 1$, $\sum_{k=1}^{c} q_{sk} = 1$. Our Bayesian model of independence for restricted pooling is

$$\begin{aligned}
\boldsymbol{p}_s|\boldsymbol{\mu}_1, \tau_1, \phi_1 &\overset{\text{iid}}{\sim} \phi_1 \text{Dirichlet}(\boldsymbol{\mu}_1\tau_1) + (1-\phi_1)\text{Dirichlet}(\mathbf{1}_p), \\
\boldsymbol{q}_s|\boldsymbol{\mu}_2, \tau_2, \phi_2 &\overset{\text{iid}}{\sim} \phi_2 \text{Dirichlet}(\boldsymbol{\mu}_2\tau_2) + (1-\phi_2)\text{Dirichlet}(\mathbf{1}_q), \\
\pi(\boldsymbol{\mu}_1, \tau_1) &= \frac{(r-1)!}{(1+\tau_1)^2}, \quad \pi(\boldsymbol{\mu}_2, \tau_2) = \frac{(c-1)!}{(1+\tau_2)^2}, \\
\phi_1 &\sim \text{Uniform}\left(\frac{1}{2}, 1\right), \quad \phi_2 \sim \text{Uniform}\left(\frac{1}{2}, 1\right),
\end{aligned}$$

where all parameters are mutually independent, and we assume that the uncertain region in the sampled data is smaller than the effectiveness. Subsequently, the joint posterior density is

$$\begin{aligned}
\pi(\boldsymbol{\Omega}|\mathbf{n}_s) = \frac{n_s! \prod_{j=1}^{r}\prod_{k=1}^{c}(p_{sj}q_{sk})^{n_{sjk}}}{\prod_{j=1}^{r}\prod_{k=1}^{c} n_{sjk}!} &\left\{ \phi_1 \frac{1}{D(\boldsymbol{\mu}_1\tau_1)} \prod_{j=1}^{r} p_{sj}^{\mu_{1j}\tau_1 - 1} + (1-\phi_1)(r-1)! \right\} \\
&\times \left\{ \phi_2 \frac{1}{D(\boldsymbol{\mu}_2\tau_2)} \prod_{k=1}^{c} p_{sk}^{\mu_{2k}\tau_2 - 1} + (1-\phi_2)(c-1)! \right\} \frac{(r-1)!}{(\tau_1+1)^2} \frac{(c-1)!}{(\tau_2+1)^2},
\end{aligned}$$

where $\boldsymbol{\Omega} = (\boldsymbol{p}_s, \boldsymbol{q}_s, \boldsymbol{\mu}_1, \tau_1, \boldsymbol{\mu}_2, \tau_2, \phi_1, \phi_2)$ and the marginal likelihood is

$$\begin{aligned}
f(\mathbf{n}_s) = \iint \frac{n_s!}{\prod_{j=1}^{r}\prod_{k=1}^{c} n_{sjk}!} \frac{(r-1)!(c-1)!}{(\tau_1+1)^2(\tau_2+1)^2} &\left\{ \frac{3}{4} \frac{D\left(\boldsymbol{\mu}_1\tau_1 + \mathbf{n}_s^{(1)}\right)}{D(\boldsymbol{\mu}_1\tau_1)} D\left(\mathbf{n}_s^{(1)} + \mathbf{1}\right) \right\} + \frac{1}{4}(r-1)! \\
&\times \left\{ \frac{3}{4} \frac{D\left(\boldsymbol{\mu}_2\tau_2 + \mathbf{n}_s^{(2)}\right)}{D(\boldsymbol{\mu}_2\tau_2)} \frac{1}{4}(c-1)!D\left(\mathbf{n}_s^{(2)} + \mathbf{1}\right) \right\} d\boldsymbol{\mu} d\boldsymbol{\tau},
\end{aligned}$$

where $\boldsymbol{\Omega} = (\boldsymbol{p}_s, \boldsymbol{q}_s, \boldsymbol{\mu}_1, \tau_1, \boldsymbol{\mu}_2, \tau_2, \phi_1, \phi_2)$, $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$, $\boldsymbol{\mu}_1 = (\mu_1, \ldots, \mu_r)$, $\boldsymbol{\mu}_2 = (\mu_1, \ldots, \mu_c)$, and $\boldsymbol{\tau} = (\tau_1, \tau_2)$. Consequently, we calculate the pooled Bayes factor based on the general model versus the independence model for restricted pooling in the s'th area. A pooled Bayes factor is

$$\mathrm{PBF}_s = \frac{(I-1)! f_1(\mathbf{n}_s)}{(r-1)!(c-1)! f_2(\mathbf{n}_s) f_3(\mathbf{n}_s)},$$

where

$$f_1(\mathbf{n}_s) = \iint \frac{1}{(\tau+1)^2} \left\{ \frac{3}{4} \frac{D(\boldsymbol{\mu}\tau + \mathbf{n}_s)}{D(\boldsymbol{\mu}\tau)} + \frac{1}{4}(I-1)! D(\mathbf{n}_s + 1) \right\} d\boldsymbol{\mu} d\tau,$$

$$f_2(\mathbf{n}_s) = \iint \frac{1}{(1+\tau_1)^2} \left\{ \frac{3}{4} \frac{D\left(\boldsymbol{\mu}_1\tau_1 + \mathbf{n}_s^{(1)}\right)}{D(\boldsymbol{\mu}_1\tau_1)} + \frac{1}{4}(r-1)! D\left(\mathbf{n}_s^{(1)} + \mathbf{1}\right) \right\} d\boldsymbol{\mu}_1 d\tau_1,$$

$$f_3(\mathbf{n}_s) = \iint \frac{1}{(1+\tau_2)^2} \left\{ \frac{3}{4} \frac{D\left(\boldsymbol{\mu}_2\tau_2 + \mathbf{n}_s^{(2)}\right)}{D(\boldsymbol{\mu}_2\tau_2)} + \frac{1}{4}(c-1)! D\left(\mathbf{n}_s^{(2)} + \mathbf{1}\right) \right\} d\boldsymbol{\mu}_2 d\tau_2.$$

3. Hierarchical Bayesian nonparameteric models

3.1. General models

In the s^{th} area, let n_{si}, $i = 1, \ldots, I(= rc)$ be the cell counts with corresponding probabilities π_{si}. For the pooled Bayes factor, the basic model is the same as parametric models. The basic model for the nonparametric version is also

$$\mathbf{n}_s | \boldsymbol{\pi}_s \overset{\text{ind}}{\sim} \mathrm{Multi}(n_s, \boldsymbol{\pi}_s), \quad s = 1, \ldots, S,$$

where $\mathbf{n}_s = (n_{s1}, \ldots, n_{sI})$ and $\boldsymbol{\pi}_s = (\pi_{s1}, \ldots, \pi_{sI})$, and $\sum_{i=1}^{I} n_{si}$. In the nonparametric version for restricted pooling, we assume that

$$\begin{aligned}
\boldsymbol{\pi}_s | G &\overset{\text{iid}}{\sim} G, \\
G &\sim DP(\alpha, G_0),\ G_0 \equiv \phi \mathrm{Dirichlet}(\boldsymbol{\mu}\tau) + (1-\phi)\mathrm{Dirichlet}(\mathbf{1}), \\
\pi(\boldsymbol{\mu}, \tau) &= \frac{(I-1)!}{(1+\tau)^2}, \\
\pi(\alpha) &= \frac{1}{(1+\alpha)^2}, \\
\phi &\sim \mathrm{Uniform}\left(\frac{1}{2}, 1\right),
\end{aligned}$$

where $\mathbf{1}$ is the $I \times 1$ vector with a value of 1. In the Dirichlet process prior, the base distribution G_0 is composed of mixed Dirichlet distributions with parameters $\boldsymbol{\mu}\tau$ and $\mathbf{1}$. Antoniak (1974). $\mathrm{DP}(\alpha, G_0)$ means Dirhcihlet process prior, where G_0 is base distribution, which can be specified as an distribution of unknown parameters defined in the space of the distribution. In this paper, this base distribution is specified as Dirichlet distribution, where α is a concentration parameter that determines the density of the distribution, i.e., the degree of parameter similarity, and the larger the value, the greater the heterogeneity of the distribution. In the non-parametric model, parameters of similar regions have the same subscript, and different characteristics have different subscripts, but if you have parameters that

are somewhat similar to other ranges, you can have more similarity in estimates, and vice versa. One of the base distributions is non-informative and reflects the uncertainty in the model. We also apply a slice-efficient sampler to fit the Dirichlet process prior. Kalli *et al.* (2011). That is, the joint posterior distribution is

$$P(\mathbf{n}, \boldsymbol{d}, \mathbf{u}|\boldsymbol{\nu}, \boldsymbol{\pi}) = \prod_{s=1}^{S} \left\{ I_{(u_s<\xi_{d_s})} \frac{w_{d_s}}{\xi_{d_s}} n_s! \prod_{i=1}^{I} \frac{\pi_{d_s i}^{n_{si}}}{n_{si}!} \right\},$$

where $w_1 = \nu_1, w_l = (1-\nu_1)\cdots(1-\nu_{l-1})\nu_l,\ \nu_l \sim B(1,\alpha),\ l = 1,\ldots,L, \xi_l = (1-\kappa)\kappa^{l-1}$. Then the joint posterior density is

$$\pi(\boldsymbol{\pi}, \boldsymbol{\nu}, \boldsymbol{d}, \mathbf{u}, \boldsymbol{\mu}, \tau, \alpha, \phi|\mathbf{n}) = \prod_{s=1}^{S} \left\{ I_{(u_s<\xi_{d_s})} \frac{w_{d_s}}{\xi_{d_s}} n_s! \prod_{i=1}^{I} \frac{\pi_{d_s i}^{n_{si}}}{n_{si}!} \right\} \prod_{l=1}^{L} \left\{ \left(\phi \frac{1}{D(\boldsymbol{\mu}\tau)} \prod_{i=1}^{I} \pi_{li}^{\mu_i \tau - 1} + (1-\phi)\frac{1}{D(\mathbf{1})} \right) \times \frac{1}{B(1,\alpha)}(1-\nu_l)^{\alpha-1} \right\} \frac{(I-1)!}{(1+\tau)^2} \frac{1}{(1+\alpha)^2}.$$

The posterior density is similar forms to the nonparametric model for adaptive pooling. However, the parameters are estimated from the data, excluding the uncertain part. Then, the marginal likelihood function is

$$f(\mathbf{n}) = \frac{\prod_{s=1}^{S} n_s!}{\prod_{s=1}^{S}\prod_{i=1}^{I} n_{si}!} \iiint \frac{(I-1)!}{(1+\tau)^2} \frac{1}{(1+\alpha)^2} \prod_{l=1}^{L} \left\{ \left(\frac{3}{4} \frac{D(\boldsymbol{\mu}\tau + \sum_{s=1}^{S} I_{(d_s=l)}\mathbf{n}_s)}{D(\boldsymbol{\mu}\tau)} + \frac{1}{4} \frac{D(\sum_{s=1}^{S} I_{(d_s=l)}\mathbf{n}_s + \mathbf{1})}{D(\mathbf{1})} \right) \frac{B(1 + \sum_{s=1}^{S} I_{(d_s=l)}, \alpha + \sum_{s=1}^{S} I_{(d_s>l)})}{B(1,\alpha)} \right\} d\alpha d\tau d\boldsymbol{\mu},$$

where $\boldsymbol{\Omega} = (\boldsymbol{\pi}, \boldsymbol{\nu}, \boldsymbol{d}, \mathbf{u}, \boldsymbol{\mu}, \tau, \alpha, \phi)$.

3.2. Independence models

In the nonparametric version, the basic model is also constructed on the same structure as the parametric model of independence. We assume that the cell counts $n_{sjk},\ j = 1,\ldots,r,\ k = 1,\ldots,c$, are the responses for the j^{th} and the k^{th} column in the sth area, $s = 1,\ldots,S$ with corresponding cell probabilities $\pi_{sjk} = p_{sj}q_{sk}$, where $p_{sj} = \sum_{k=1}^{c} \pi_{sjk}$ and $q_{sk} = \sum_{j=1}^{r} \pi_{sjk}$. For our pooled Bayes factor of independence, we propose the following:

$$\mathbf{n}_s|\boldsymbol{p}_s, \boldsymbol{q}_s \overset{\text{ind}}{\sim} \text{Multi}\left(n_s, \text{vec}\left(\boldsymbol{p}_s', \boldsymbol{q}_s'\right)\right), \quad s = 1,\ldots,S,$$

where $\mathbf{n}_s = (n_{s11},\ldots,n_{s1c},\ldots,n_{sr1},\ldots,n_{src})$, $n_s = \sum_{j=1}^{r}\sum_{k=1}^{c} n_{sjk}$, $\boldsymbol{p}_s = (p_{s1},\ldots,p_{sr})$,$\boldsymbol{q} = (q_{s1},\ldots,q_{sc})$, $\sum_{j-1}^{r} p_{sj} = 1$, $\sum_{k=1}^{c} q_{sk} = 1$. We construct the nonparametric version of the restricted pooling model, which is built using the mixture Dirichlet distribution to reflect the uncertainty. Our nonparametric model for the restricted pooling of independence is

$$\boldsymbol{p}_s|G_1 \overset{\text{iid}}{\sim} G_1, \quad G_1 \sim DP(\alpha_1, G_{01}), \quad G_{01} \equiv \phi_1 \text{Dirichlet}(\boldsymbol{\mu}_1, \tau_1) + (1-\phi_1)\text{Dirichlet}(\mathbf{1}_p),$$

$$\boldsymbol{q}_s|G_2 \overset{\text{iid}}{\sim} G_2, \quad G_2 \sim DP(\alpha_1, G_{02}), \quad G_{02} \equiv \phi_2 \text{Dirichlet}(\boldsymbol{\mu}_2, \tau_2) + (1-\phi_2)\text{Dirichlet}(\mathbf{1}_q),$$

$$\pi(\boldsymbol{\mu}_1, \tau_1) = \frac{(r-1)!}{(\tau_1+1)^2}, \quad \pi(\alpha_1) = \frac{1}{(\alpha_1+1)^2}, \quad \pi(\boldsymbol{\mu}_2, \tau_2) = \frac{(c-1)!}{(\tau_2+1)^2}, \quad \pi(\alpha_2) = \frac{1}{(\alpha_2+1)^2},$$

$$\phi_1 \sim \text{Uniform}\left(\frac{1}{2}, 1\right), \quad \phi_2 \sim \text{Uniform}\left(\frac{1}{2}, 1\right), \quad \tau_1, \tau_2 > 0, \quad \alpha_1, \alpha 2 > 0,$$

where $\mathbf{1}_p$ is $r\times 1$ vector with a component equal to 1, $\mathbf{1}_q$ is $c\times 1$ vector, and the Dirichlet processes are composed of the base distributions with a mixture Dirichlet distribution. The uncertainty is reflected in the Dirichlet distribution with parameter vectors $\mathbf{1}_p$ and $\mathbf{1}_q$. Then, the joint posterior density based on the above model structure is

$$\begin{aligned}\pi(\boldsymbol{\Omega}|\mathbf{n}) = &\prod_{s=1}^{S}\left[\left\{I_{(u_{1s}<\xi_{d_{1s}})}\frac{w_{d_{1s}}}{\xi_{d_{1s}}}I_{(u_{2s}<\xi_{d_{2s}})}\frac{w_{d_{2s}}}{\xi_{d_{2s}}}n_s!\prod_{j=1}^{r}\prod_{k=1}^{c}\frac{(p_{d_{1s}j}q_{d_{2s}k})^{n_{sjk}}}{n_{sjk}!}\right\}\right]\\
&\times\prod_{l_1=1}^{L_1}\left[\left\{\phi_1\frac{1}{D(\boldsymbol{\mu}_1\tau_1)}\prod_{j=1}^{r}p_{1j}^{\mu_{1j}\tau_1}+(1-\phi_1)(r-1)!\right\}\frac{1}{B(1,\alpha_1)}(1-\nu_{1l})^{\alpha_1-1}\right]\\
&\times\prod_{l_2=1}^{L_2}\left[\left\{\phi_2\frac{1}{D(\boldsymbol{\mu}_2\tau_2)}\prod_{k=1}^{c}q_{2k}^{\mu_{2k}\tau_2}+(1-\phi_2)(c-1)!\right\}\frac{1}{B(1,\alpha_2)}(1-\nu_{2l})^{\alpha_2-1}\right]\\
&\times\frac{(c-1)!}{(\tau_2+1)^2}\frac{1}{(\alpha_2+1)^2}\frac{(r-1)!}{(\tau_1+1)^2}\frac{1}{(\alpha_1+1)^2},\end{aligned}$$

where $\boldsymbol{\Omega}=(\boldsymbol{p},\boldsymbol{q},\boldsymbol{\nu}_1,\boldsymbol{\nu}_2,\boldsymbol{d}_1,\boldsymbol{d}_2,\mathbf{u}_1,\mathbf{u}_2,\boldsymbol{\mu}_1,\tau_1,\alpha_1,\boldsymbol{\mu}_2,\tau_2,\alpha_2,\phi_1,\phi_2)$, L_1 is largest number of cluster for parameters $\boldsymbol{p}_s$ for $s=1,\ldots,S$ and L_2 is largest number of cluster for parameters $\boldsymbol{q}_s$ for $s=1,\ldots,S$, and we assume that the structure of cluster between $\boldsymbol{p}_s$ and $\boldsymbol{q}_s$ is independent. Our marginal likelihood of independence is

$$\begin{aligned}f(\mathbf{n}) = &\sum_{s=1}^{S}\sum_{d_{1s}=1}^{L_1}\sum_{d_{2s}=1}^{L_2}\left[\frac{\prod_{s=1}^{S}n_s!}{\prod_{s=1}^{S}\prod_{i=1}^{I}n_{si}!}\right]\int\frac{(r-1)!}{(\tau_1+1)^2}\frac{1}{(\alpha_1+1)^2}\frac{(c-1)!}{(\tau_2+1)^2}\frac{1}{(\alpha_2+1)^2}\\
&\times\prod_{l_1=1}^{L_1}\left[\left\{\frac{3}{4}\frac{D\left(\boldsymbol{\mu}_1\tau_1+\sum_{s=1}^{S}I_{(d_{1s}=l_1)}\mathbf{n}_s^{(1)}\right)}{D(\boldsymbol{\mu}_1\tau_1)}+\frac{1}{4}(r-1)!D\left(\sum_{s=1}^{S}I_{(d_{1s}=l_1)}\mathbf{n}_s^{(1)}+\mathbf{1}\right)\right\}\right]\\
&\times\prod_{l_2=1}^{L_2}\left[\left\{\frac{3}{4}\frac{D\left(\boldsymbol{\mu}_2\tau_2+\sum_{s=1}^{S}I_{(d_{2s}=l_2)}\mathbf{n}_s^{(2)}\right)}{D(\boldsymbol{\mu}_2\tau_2)}+\frac{1}{4}(c-1)!D\left(\sum_{s=1}^{S}I_{(d_{2s}=l_2)}\mathbf{n}_s^{(2)}+\mathbf{1}\right)\right\}\right]\\
&\times\prod_{l_1=1}^{L_1}\frac{B\left(1+\sum_{s=1}^{S}I_{(d_{1s}=l_1)},\alpha_1+\sum_{s=1}^{S}I_{(d_{1s}>l_1)}\right)}{B(1,\alpha_1)}\times\prod_{l_2=1}^{L_2}\frac{B\left(1+\sum_{s=1}^{S}I_{(d_{2s}=l_2)},\alpha_2+\sum_{s=1}^{S}I_{(d_{2s}>l_2)}\right)}{B(1,\alpha_2)}d\boldsymbol{\Omega}',\end{aligned}$$

where $\boldsymbol{\Omega}'=(\boldsymbol{d}_1,\boldsymbol{d}_2,\boldsymbol{\mu}_1,\tau_1,\alpha_1,\boldsymbol{\mu}_2,\tau_2,\alpha_2)$ and the pooled Bayes factor of the general model versus the independence model for the s'th area is

$$\mathrm{PBF}_s=\frac{(I-1)!f_1(\mathbf{n})}{(r-1)!(c-1)!f_2(\mathbf{n})},$$

where

$$\begin{aligned}f_1(\mathbf{n}) = &\frac{\prod_{s=1}^{S}n_s!}{\prod_{s=1}^{S}\prod_{i=1}^{I}n_{si}!}\int\frac{(I-1)!}{(1+\tau)^2}\frac{1}{(1+\alpha)^2}\prod_{l=1}^{L}\left\{\left(\frac{3}{4}\frac{D\left(\boldsymbol{\mu}\tau+\sum_{s=1}^{S}I_{(d_s=l)}\mathbf{n}_s\right)}{D(\boldsymbol{\mu}\tau)}\right.\right.\\
&\left.\left.\times\frac{1}{4}\frac{D\left(\sum_{s=1}^{S}I_{(d_s=l)}\mathbf{n}_s+\mathbf{1}\right)}{D(\mathbf{1})}\right)\frac{B(1+\sum_{s=1}^{S}I_{(d_s=l)},\alpha+\sum_{s=1}^{S}I_{(d_s>l)})}{B(1,\alpha)}\right\}d\boldsymbol{\Omega}\end{aligned}$$

Table 1: BMI and BMD contingency tables from the NHANES III

Area	N	BMI											
		Underweight			Optimal			Overweight			Obese		
		BMD			BMD			BMD			BMD		
		N	OPE	OP	N	OPE	OP	N	OPE	OP	N	OPE	OP
1	78	5	0	0	11	8	4	12	9	1	27	1	0
2	65	1	2	1	15	5	4	15	1	1	18	2	0
3	73	2	3	2	17	7	3	18	6	1	10	3	1
4	71	5	5	0	15	6	3	15	6	2	13	1	0
5	66	5	1	2	16	11	3	4	7	4	8	5	0
6	93	2	1	1	18	6	3	17	9	1	32	3	0
7	106	2	2	1	15	7	4	35	7	2	26	5	0
8	161	4	3	1	39	15	3	29	15	3	38	11	0
9	266	12	10	2	45	16	4	59	19	5	80	10	4
10	90	7	2	1	18	17	4	11	12	0	10	8	0
11	81	4	1	1	15	11	2	22	4	1	18	2	0
12	408	11	7	2	79	32	8	108	36	6	98	20	1
13	62	1	1	2	16	6	1	5	6	0	20	3	1
14	104	7	0	1	22	3	0	28	6	3	31	3	0
15	153	5	6	4	34	21	8	33	9	2	27	4	0
16	202	10	4	2	40	17	7	43	15	0	56	8	0
17	59	1	1	1	19	8	2	5	4	1	15	2	0
18	95	2	1	1	24	8	1	24	6	1	23	4	0
19	73	2	6	2	10	9	3	12	5	3	16	5	0
20	79	2	2	2	14	9	2	15	7	1	17	7	1
21	65	0	0	1	16	5	8	8	7	1	16	3	0
22	59	1	2	2	6	8	4	18	4	0	13	1	0
23	116	1	4	3	29	8	5	20	12	3	24	6	1
24	65	1	3	1	13	6	3	11	2	1	22	1	1
25	54	2	2	0	13	7	1	9	1	1	13	4	1
26	59	5	5	0	18	3	2	11	6	0	5	4	0
27	94	4	2	0	9	7	2	23	11	2	27	7	0
28	93	4	3	0	23	3	2	14	5	1	35	3	0
29	76	1	1	2	18	6	5	17	6	4	13	3	0
30	115	2	0	0	24	7	4	31	3	2	39	3	0
31	94	6	2	3	26	7	1	21	4	0	24	0	0

Note. N = Normal, OPE = Osteopenia, OP = Osteoporosis.

and

$$
\begin{aligned}
f_2(\mathbf{n}) = \sum_{s=1}^{S} \sum_{d_{1s}=1}^{L_1} \sum_{d_{2s}=1}^{L_2} \left[\frac{\prod_{s=1}^{S} n_s!}{\prod_{s=1}^{S} \prod_{i=1}^{I} n_{si}!} \right] \int \frac{(r-1)!}{(\tau_1+1)^2} \frac{1}{(\alpha_1+1)^2} \frac{(c-1)!}{(\tau_2+1)^2} \frac{1}{(\alpha_2+1)^2} \\
\times \prod_{l_1=1}^{L_1} \left[\left\{ \frac{3}{4} \frac{D\left(\boldsymbol{\mu}_1\tau_1 + \sum_{s=1}^{S} I_{(d_{1s}=l_1)} \mathbf{n}_s^{(1)}\right)}{D\left(\boldsymbol{\mu}_1\tau_1\right)} + \frac{1}{4}(r-1)! D\left(\sum_{s=1}^{S} I_{(d_{1s}=l_1)} \mathbf{n}_s^{(1)} + \mathbf{1} \right) \right\} \right] \\
\times \prod_{l_2=1}^{L_2} \left[\left\{ \frac{3}{4} \frac{D\left(\boldsymbol{\mu}_2\tau_2 + \sum_{s=1}^{S} I_{(d_{2s}=l_2)} \mathbf{n}_s^{(2)}\right)}{D\left(\boldsymbol{\mu}_2\tau_2\right)} + \frac{1}{4}(c-1)! D\left(\sum_{s=1}^{S} I_{(d_{2s}=l_2)} \mathbf{n}_s^{(2)} + \mathbf{1} \right) \right\} \right] \\
\times \prod_{l_1=1}^{L_1} \frac{B\left(1+\sum_{s=1}^{S} I_{(d_{1s}=l_1)}, \alpha_1 + \sum_{s=1}^{S} I_{(d_{1s}>l_1)}\right)}{B(1,\alpha_1)} \times \prod_{l_2=1}^{L_2} \frac{B\left(1+\sum_{s=1}^{S} I_{(d_{2s}=l_2)}, \alpha_2 + \sum_{s=1}^{S} I_{(d_{2s}>l_2)}\right)}{B(1,\alpha_2)} d\boldsymbol{\Omega}',
\end{aligned}
$$

where $\boldsymbol{\Omega} = (\alpha, \tau, \boldsymbol{\mu})$ and $\boldsymbol{\Omega}' = (\boldsymbol{d}_1, \boldsymbol{d}_2, \boldsymbol{\mu}_1, \tau_1, \alpha_1, \boldsymbol{\mu}_2, \tau_2, \alpha_2)$.

Table 2: Posterior summaries of pooling parameter (μ, τ) in parametric model

	PM	PSD	NSE	95%CI
1. General model				
μ_1	0.078	0.046	0.001	(0.013, 0.193)
μ_2	0.065	0.034	0.000	(0.012, 0.143)
μ_3	0.049	0.028	0.000	(0.008, 0.114)
μ_4	0.147	0.075	0.002	(0.024, 0.309)
μ_5	0.094	0.051	0.001	(0.02, 0.214)
μ_6	0.066	0.041	0.001	(0.013, 0.174)
μ_7	0.124	0.062	0.002	(0.023, 0.259)
μ_8	0.086	0.049	0.001	(0.017, 0.204)
μ_9	0.049	0.034	0.001	(0.008, 0.138)
μ_{10}	0.138	0.069	0.002	(0.026, 0.287)
μ_{11}	0.074	0.045	0.001	(0.017, 0.188)
μ_{12}	0.029	0.036	0.000	(0.002, 0.141)
τ	21.062	20.994	0.140	(0.354, 71.639)
2. Independence model				
2.1. Row table				
μ_1	0.620	0.086	0.001	(0.422, 0.743)
μ_2	0.262	0.061	0.001	(0.155, 0.388)
μ_3	0.118	0.063	0.001	(0.044, 0.262)
τ	30.012	38.395	0.273	(2.059, 146.248)
2.2. Column table				
μ_1	0.103	0.058	0.000	(0.028, 0.249)
μ_2	0.326	0.066	0.001	(0.180, 0.442)
μ_3	0.285	0.061	0.001	(0.162, 0.403)
μ_4	0.287	0.076	0.000	(0.145, 0.434)
τ	72.171	66.117	0.741	(1.359, 254.275)

Note. PM = Posterior mean, PSD = Posterior standard deviation, NSE = Numerical standard error(are obtained from the batch means method).

4. Data analysis

As medical technology improves, the purpose of clinical care is not merely survival but to bring a change in terms of improving the quality of life. Therefore, determining the association among different clinical factors and predicting diseases based on other clinical factors have become the most important issues in the clinical field. As part of these changes, the National Center for Health Statistics has implemented a survey program, the NHANES, since the early 1960s in the U.S. The NHANES is a major survey program to measure and assess the health and nutritional status of people. It combines physical examination and interview data collected by mobile examination centers in the U.S. As indicated in Table 1, the BMI and BMD data are also a part of this survey program, and the data were collected between 1994 and 1998. We apply our method to the two-way contingency tables with the BMI and BMD category data from NHANES III to compare and assess our test of independence with the pooled Bayes factor.

Although NHANES III is a large-scale survey, the cell frequency of the contingency tables for identifying the characteristics of small regions was often less than 5. In general, it cannot be used for the independence test using the chi-square test as in the case above, and the limitation is that the reliability of the test results is very low. On the other hand, in the case of the pooled Bayes factor, which has a structure that can provide additional information on cells below the value of 5 by borrowing information from surrounding regions, it is modeled by directly reflecting the characteristics of the data and a reliable test can be performed through this. A summary of the estimated pooling parameter is presented in Table 2 and 3. The estimated summaries of μ and τ, which are pooling parameters es-

Table 3: Posterior summaries of pooling parameter (μ, τ) in nonparametric model

	PM	PSD	NSE	95%CI
1. General model				
μ_1	0.055	0.012	0.0004	(0.036, 0.081)
μ_2	0.042	0.010	0.0004	(0.025, 0.062)
μ_3	0.028	0.007	0.0005	(0.015, 0.044)
μ_4	0.171	0.017	0.0024	(0.138, 0.203)
μ_5	0.102	0.015	0.0014	(0.074, 0.132)
μ_6	0.052	0.011	0.0007	(0.031, 0.074)
μ_7	0.167	0.016	0.0011	(0.137, 0.200)
μ_8	0.092	0.013	0.0008	(0.067, 0.118)
μ_9	0.032	0.009	0.0003	(0.016, 0.052)
μ_{10}	0.181	0.016	0.0015	(0.147, 0.209)
μ_{11}	0.063	0.011	0.0003	(0.043, 0.084)
μ_{12}	0.016	0.005	0.0002	(0.008, 0.026)
τ	158.788	12.363	0.1750	(124.2, 170.2)
2. Independence model				
2.1. Row table				
μ_1	0.619	0.023	0.0020	(0.575, 0.663)
μ_2	0.268	0.020	0.0006	(0.232, 0.309)
μ_3	0.113	0.016	0.0016	(0.082, 0.143)
τ	139.029	27.114	2.6115	(68.1, 169.5)
2.2. Column table				
μ_1	0.063	0.010	0.0004	(0.045, 0.084)
μ_2	0.308	0.018	0.0011	(0.273, 0.345)
μ_3	0.280	0.017	0.0008	(0.248, 0.313)
μ_4	0.350	0.020	0.0009	(0.311, 0.390)
τ	159.462	10.916	0.3579	(129.7, 170.1)

Note. PM = Posterior mean, PSD = Posterior standard deviation, NSE = Numerical standard error(are obtained from the batch means method).

timated in the general and the independence model, show that the posterior standard deviation is very low with a maximum of 0.086. In particular, the stability of the chain was confirmed, given that the numerical standard error calculated through the batch means method is estimated to be below 0.0002 in most parameters.

In Table 4, we can compare the chi-square test results by region and the results of the pooled Bayes factor (PBF) calculated through the parametric and nonparametric Bayesian model. If the $2 * \log(\text{PBF})$ is in (0, 1) or the p-value is in (0.05, 0.10), we obtain borderline evidence against the null hypothesis. If the log(PBF) is in (1, 3) or the p-value is in (0.025, 0.05), we obtain positive evidence against the null hypothesis. If the log(PBF) is in (3, 5) or the p-value is in (0.01, 0.025), we derive strong evidence against the null hypothesis. If the log(BF) is greater than 5 or the p-value is in (0.000, 0.010), we derive very strong evidence against the null hypothesis (refer to Kass and Raftery, 1995). Although there is no evidence that the relationship between regional BMI and BMD differs clinically, the chi-square test results indicate contradictory results by region. Although the pooled Bayes factor test results show differences in intensity by region, there was in general an association between BMI and BMD. When the chi-square test concludes that the two factors between regions are independent, the frequency of 5 or less is widely distributed, and it is difficult to confirm the reliability of the test. However, Jo *et al.* (2021) found consistent statistical test results when tests were conducted using the pooled Bayes factor calculated with the restricted pooling in which a pooling effect between regions was confirmed. In particular, areas 1, 15, 19, and 22 had uncertainties classified by ϕ. While the test results indicating the association between BMI and BMD are similar, the intensity of the test is

Table 4: Comparisons of the Pearson's Chi-Squared test and pooled Bayes tests

Area	n	Chi-Squared	$\log(\text{PBF}_p)$	$\log(\text{PBF}_{np})$
1	78	0.001	2.966	3.082$^{£}$
2	65	0.055$^{\$}$	2.648	2.897
3	73	0.396$^{\$}$	2.377	2.842
4	71	0.178$^{\$}$	2.562	2.898
5	66	0.192$^{\$}$	2.676	2.949
6	93	0.034	2.618	2.866
7	106	0.080$^{\$}$	2.526	2.822
8	161	0.387$^{\$}$	2.115	2.595
9	266	0.022	2.416	2.786
10	90	0.333$^{\$}$	2.471	2.834
11	81	0.079$^{\$}$	2.529	2.859
12	408	0.033	2.216	2.648
13	62	0.002	2.716	2.923
14	104	0.239$^{\$}$	2.245	2.491
15	153	0.002	2.89	3.166$^{£}$
16	202	0.002	2.764	2.984
17	59	0.191$^{\$}$	2.549	2.829
18	95	0.216$^{\$}$	2.130	2.578
19	73	0.080$^{\$}$	2.748	3.046$^{£}$
20	79	0.287$^{\$}$	2.282	2.769
21	65	0.006	2.825	2.990
22	59	0.001	3.030$^{£}$	3.184$^{£}$
23	116	0.013	2.657	2.981
24	65	0.030	2.702	2.953
25	54	0.706$^{\$}$	2.272	2.667
26	59	0.208$^{\$}$	2.579	2.838
27	94	0.325$^{\$}$	2.368	2.700
28	93	0.092$^{\$}$	2.468	2.694
29	76	0.257$^{\$}$	2.495	2.857
30	115	0.131$^{\$}$	2.354	2.532
31	94	0.001	2.759	2.936

$^{\$}$ Areas with independence of BMI and BMD; $^{£}$ Areas with more strong dependence of BMI an BMD.

more robust, which implies that the BMI and BMD association of the regions is higher than in other regions.

5. Concluding remarks

When the number of data is small in the small-region estimation, we present the calculated results by theoretically developing the pooled Bayes factor based on the restricted pooling model, a model proposed by Jo *et al.* (2021) that borrows information from similar nearby areas and performs a pooled parameter estimation, and proposed the Bayes test that performs a statistical test using the calculated PBF. The performance of parameter estimation was found to have an increased precision by borrowing nearby information, reducing the number of parameters and increasing the number of data for parameter estimation, as proposed by Jo *et al.* (2021). In this case, the pooling of excess data may excessively reduce fluctuations caused by the local effect and cause an over-shrinkage problem in the parameter estimation. To prevent this, a restricted pooling model was proposed that limits regions with different characteristics, and it was confirmed that the performance of the proposed model was better than the pooling of all existing data. Using the restricted pooling model, we proposed the Bayes test,

which borrows information from similar regions and performs the independence test and were able to produce consistent test results for all regions, unlike the existing chi-square test. In the restricted pooling model, compared to the parametric model, which indirectly pools local information through the same hyper-parameter, in case of a nonparametric model that directly pools using parameters with the same subscript in a similar area. Because the test statistics are calculated using only the separate regional information in specific areas with different characteristics, the smoothing effect of the parameter is reduced compared to the parametric model. Therefore, the regional specific test results can be shown, and even in the data analysis it can be seen that there are many areas with different results in the nonparametric model. In other words, in the case of a parametric model, since the estimation of hyper-parametric may include indirect information from other regions even in different regional characteristics, similar test statistics can be calculated, and study over the possibility of over-shrinkage should be considered. In particular, given that the number of clusters calculated through the nonparametric Bayesian model was 1 to 3, the distribution of BMD and BMI did not differ significantly by region. Since information about regions with specific characteristics is classified as regions with uncertainty through the parameter ϕ, it was also possible to confirm regions with different test results' intensity.

References

Antoniak CE (1974). Mixture of Dirichlet processes with applications to Bayesian nonparametric problems, *The Annals of Statistics*, **2**, 1152–1174.

Consonni G and Veronese P (1995). A Bayesian method for combining results from several binomial experiments, *Journal of the American Statistical Association*, **90**, 935–944.

DuMouchel WH and Harris JE (1983). Bayes methods for combining the results of cancer studies in humans and other species, *Journal of the American Statistical Association*, **78**, 293–308.

Dunson DB (2009). Nonparametric Bayes local partition models for random effects, *Biometrika*, **96**, 249–262.

Dunson DB, Xue Y, and Carin L (2009). The matrix stick-breaking process: flexible Bayes meta-analysis, *Journal of the American Statistical Association*, **103**, 317–327.

Evans R and Sedransk J (2003). Bayesian methodology for combining the results from different experiments when the specifications for pooling are uncertain: II, *Journal of Statiatical Planning and Inference*, **111**, 95–100.

Ferguson TS (1973). A Bayesian analysis of some nonparametric problems, *The Annals of Statistics*, **1**, 209–230.

Jo A, Nandram B, and Kim DH (2021). Bayesian pooling for analyzing categorical data from small areas, *Survey Methodology*, **1**, 191–213.

Kalli M, Griffin JE, and Walker SG (2011). Slice sampling mixture models, *Statistics and Computing*, **21**, 93–105.

Malec D and Sedransk J (1992). Bayesian methodology for combining the results from different experiments when the specifications for pooling are uncertain, *Biometrika*, **79**, 593–601.

Nandram B (1998). A Bayesian analysis of the three-stage hierarchical multinomial model, *Journal of Statistical Computation and Simulation*, **61**, 97–112.

Nandram B and Kim H (2002). Marginal likelihood for a class of Bayesian generalized linear models, *Journal of Statistical Computation and Simulation*, **72**, 319–340.

Nandram B, Zhou J, and Kim DH (2019). A pooled Bayes test of independence for sparse contingency tables from small area, *Journal of Statistical Computation and Simulation*, **89**, 899–926.

Walker SG (2007). Sampling the Dirichlet mixture model with slices, *Communications in Statistics-Simulation and Computation*, **36**, 45–54.

Received January 06, 2022; Revised May 5, 2022; Accepted June 15, 2022