Web-based Supplementary Materials for "Bayesian model selection in order-restricted two-way ANOVA mixed models"

Yonggang Ji, Haifang Shi

School of Science, Civil Aviation University of China, Tianjin, China

ygji@cauc.edu.cn;hfshi@cauc.edu.cn

1. Sensitivity of the SSVS in real data analysis

As George and McCulloch(1993) have proved that there are only small fluctuations to the mixture variance parameters $c\phi_i$ in SSVS by simulations, we first fixed $\phi_i = 1$ following the suggestion of Walli and Wagner (2011), and simply considered four priors with different values of c to check the sensitivity of SSVS method in real data analysis. The posterior probabilities of all the possible models and the posterior estimates for the parameters are listed in Table 1 and Table 2, respectively. The results of Table 1 show that SSVS with $c = 1/10, 1/100$, NMIG, Kuo&Mallick tend to choose the same model, whereas SSVS with $c = 1/1000, 1/10000$ and Shang tend to choose the same model. SSVS method does not select the same model under different values of c, but SSVS always choose the same top two models. It can be observed from Table 2 that there is no significant difference in the performance of the posterior estimates with different values of c. All mentioned above indicate that the results of model selection are generally not sensitive to the parameters $c\phi_i$ in SSVS. Moreover, We further did a more careful check to the top two models selected by SSVS under different values of

c. The difference between the posterior probabilities of the top two models are list in 3. We noticed that Shang and SSVS have similar performance. The small gap between the values of the posterior probabilities of H_1 and H_F, suggests that these two models can be quite difficult to distinguish using SSVS and Shang, though the effect of δ_1 is significant in H_F. This interesting observation motivates us to think more deeply about the relationship between the SSVS and the Shang method, which we plan to study in our future research.

Table 1: Results of the posterior probabilities of all the possible models for different priors in blood lead level data.

Model	*Shang*	*NMIG*	*Kuo&Mallick*	SSVS			
				$c = 1/10$	$c = 1/100$	$c = 1/1000$	$c = 1/10000$
H_0	0.0535	0.0000	0.0000	0.0033	0.0000	0.0000	0.0000
H_1	0.1890	**0.3732**	**0.3997**	**0.2962**	**0.3655**	0.3740	0.3722
H_2	0.1898	0.2648	0.2555	0.2146	0.1950	0.1866	0.1764
H_3	0.0584	0.0011	0.0019	0.0173	0.0057	0.0024	0.0031
H_{12}	0.1846	0.2373	0.1955	0.2148	0.0754	0.0416	0.0390
H_{13}	0.0633	0.0037	0.0052	0.0186	0.0036	0.0015	0.0016
H_{23}	0.0478	0.0000	0.0001	0.0061	0.0001	0.0000	0.0000
H_F	**0.2136**	0.1199	0.1421	0.2290	0.3547	**0.3939**	**0.4077**

Table 2: Posterior means (mean), variances (SD) and 95% credible intervals (CrI) for blood lead level data.

Mthodes		μ_1	σ_τ^2	σ^2	δ_1	δ_2	δ_3
Shang	Mean	23.535	25.424	6.217	0.546	0.507	1.362
	Var	1.266	32.463	1.278	0.836	0.654	1.289
	CrI	(21.01,25.46)	(16.47,38.64)	(4.66,8.87)	(0,3.01)	(0,2.64)	(0,3.53)
NMIG	Mean	23.676	25.612	5.57	0.287	0.41	1.748
	Var	0.615	31.941	0.430	0.159	0.221	0.239
	CrI	(22.12,25.2)	(16.67,38.6)	(4.43,7)	(0,1.3)	(0,1.46)	(0.75,2.66)
Kuo&Mallick	Mean	23.672	25.599	5.571	0.278	0.438	1.732
	Var	0.621	32.431	0.431	0.173	0.240	0.251
	CrI	(22.1,25.21)	(16.62,38.84)	(4.43,7)	(0,1.3)	(0,1.48)	(0.72,2.66)
SSVS $c\phi_i = 1/10$	Mean	23.586	25.639	5.534	0.426	0.555	1.403
	Var	0.589	32.722	0.424	0.091	0.124	0.194
	CrI	(22.06,25.08)	(16.67,38.94)	(4.41,6.95)	(0.02,1.16)	(0.04,1.35)	(0.46,2.22)
SSVS $c = 1/100$	Mean	23.618	25.625	5.548	0.368	0.562	1.435
	Var	0.602	32.595	0.428	0.129	0.177	0.195
	CrI	(22.08,25.14)	(16.76,38.94)	(4.41,6.97)	(0.01,1.22)	(0.02,1.42)	(0.53,2.26)
SSVS $c = 1/1000$	Mean	23.624	25.621	5.55	0.356	0.579	1.433
	Var	0.6	32.434	0.426	0.148	0.192	0.187
	CrI	(22.1,25.14)	(16.68,38.79)	(4.42,6.97)	(0,1.24)	(0.01,1.44)	(0.57,2.27)
SSVS $c = 1/10000$	Mean	23.616	25.619	5.553	0.35	0.585	1.431
	Var	0.607	32.616	0.427	0.152	0.195	0.189
	CrI	(22.07,25.14)	(16.65,38.93)	(4.41,6.97)	(0,1.24)	(0,1.45)	(0.564,2.26)

Table 3: The difference between the posterior probabilities of the top two models selected by SSVS

Model	*Shang*	*NMIG*	*Kuo&Mallick*	SSVS			
				$c = 1/10$	$c = 1/100$	$c = 1/1000$	$c = 1/10000$
H_1	0.1890	0.3732	0.3997	0.2962	0.3655	0.3740	0.3722
H_F	0.2136	0.1199	0.1421	0.2290	0.3547	0.3939	0.4077
difference	-0.0246	0.2533	0.2576	0.0672	0.0108	-0.0199	-0.0355

Table 4: Results of the posterior probabilities of all the possible models for different priors in blood lead level data.

Model	Posterior prob					
	$c\phi_i = 1/10$	$c\phi_i = 1/100$	$c\phi_i = 1/500$	$c\phi_i = 1/1000$	$c\phi_i = 5000$	$c\phi_i = 1/10000$
H_0	0.0033	0.0000	0.0000	0.0000	0.0000	0.0000
H_1	**0.2962**	**0.3655**	0.3635	0.3740	0.3606	0.3722
H_2	0.2146	0.1950	0.1847	0.1866	0.1860	0.1764
H_3	0.0173	0.0057	0.0037	0.0024	0.0028	0.0031
H_{12}	0.2148	0.0754	0.0495	0.0416	0.0337	0.0390
H_{13}	0.0186	0.0036	0.0013	0.0015	0.0027	0.0016
H_{23}	0.0061	0.0001	0.0000	0.0000	0.0003	0.0000
H_F	0.2290	0.3547	**0.3973**	**0.3939**	**0.4140**	**0.4077**

References

George, E. L., McCulloch, R. E., 1993. Variable selection via Gibbs sampling. Journal of the American Statistical Association 88, 881C889.

Walli, G. M., Wagner, H., 2011. Comparing spike and slab priors for bayesian variable selection. Austrian Journal oF Statistics 40, 241C264.