STA 6352, Report 7.3

Carson Slater *Baylor University*

Problem

Exchangeable prior distributions: suppose it is known a priori that the $2J$ parameters $\theta_1, \ldots, \theta_{2J}$ are clustered into two groups, with exactly half being drawn from a $\mathcal{N}(1, 1)$ distribution, and the other half being drawn from a $\mathcal{N}(-1, 1)$ distribution, but we have not observed which parameters come from which distribution.

(a)

Are $\theta_1, \ldots, \theta_{2J}$ exchangeable under this prior distribution?

$\theta_1, \ldots, \theta_{2J}$ are in fact exchangeable under this prior, because you can characterize their joint distribution. Their exchangeable joint probability distribution would be the product of the $2J$ independent normals averaged over all possible permutations.

$$p(\boldsymbol{\theta}) = \binom{2J}{J}^{-1} \sum_p \left(\prod_{j=1}^{J} \mathcal{N}(\theta_{pj}|1, 1) \prod_{j=J+1}^{2J} \mathcal{N}(\theta_{pj}| - 1, 1) \right),$$

where the sum is over all permutations p of $(1, \ldots, 2J)$. This density is invariant to permutations of the indices.

(b)

Show that this distribution cannot be written as a mixture of independent and iden- tically distributed components.

Suppose that we define $\phi_1, \ldots, \phi_{2J}$, where half of the ϕ_j's are 1 and the other half are 1. Then setting $\theta_j|\phi_j \sim \mathcal{N}(\phi_j, 1)$. Then it is easy to show first that $\text{Cov}(\phi_i, \phi_j) < 0$ and then that $\text{Cov}(\theta_i, \theta_j) < 0$. We can show this analytically but we elect to show this computationally.

Then, by Exercise 5.5, $p(\theta_1, \ldots, \theta_{2J})$ cannot be written as a mixture of iid components.

(c)

Why can we not simply take the limit as $J \to \infty$ and get a counterexample to de Finetti's theorem?

In the limit as $J \to \infty$, the negative correlation between θ_i and θ_j approaches zero, and the joint distribution approaches iid. To put it another way, as $J \to \infty$, the distinction disappears between (1) independently

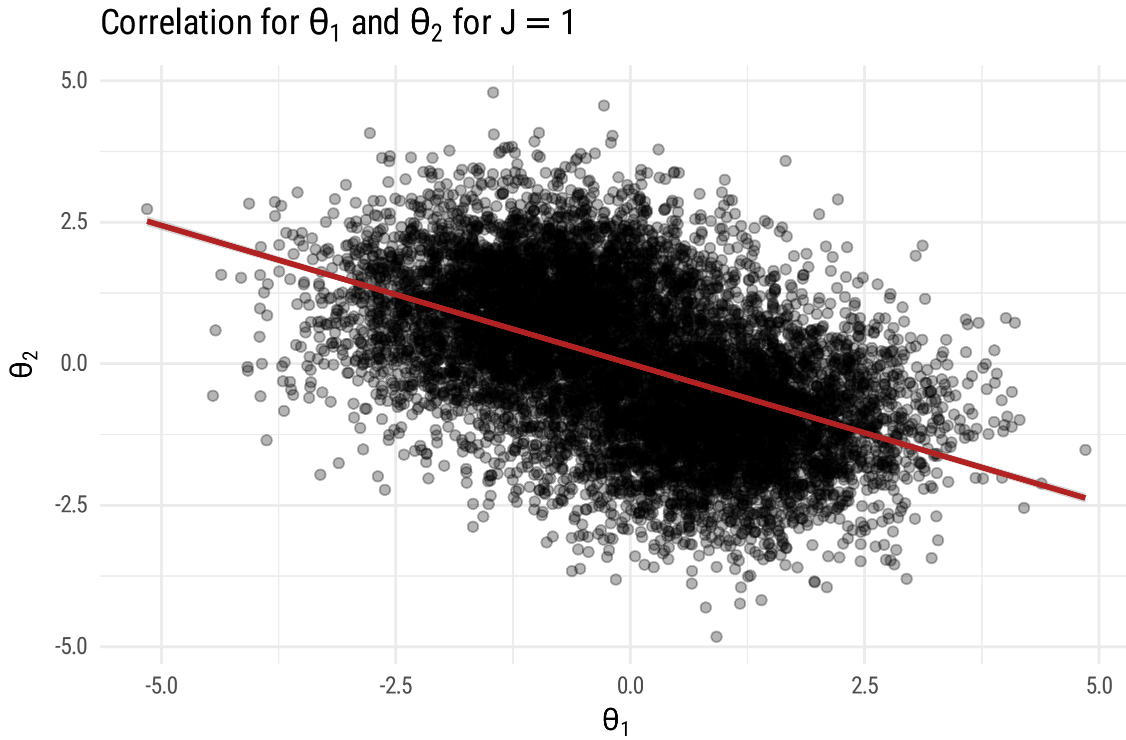

Figure 1: Clearly there is a negative relationship between these variables.

assigning each θ_j to one of two groups, and (2) picking exactly half of the θ_j's for each group. We demonstrate this via a simulation. As J grows larger in figure 1, we see that the sample covariance converges. to around zero; however it is generally negative.

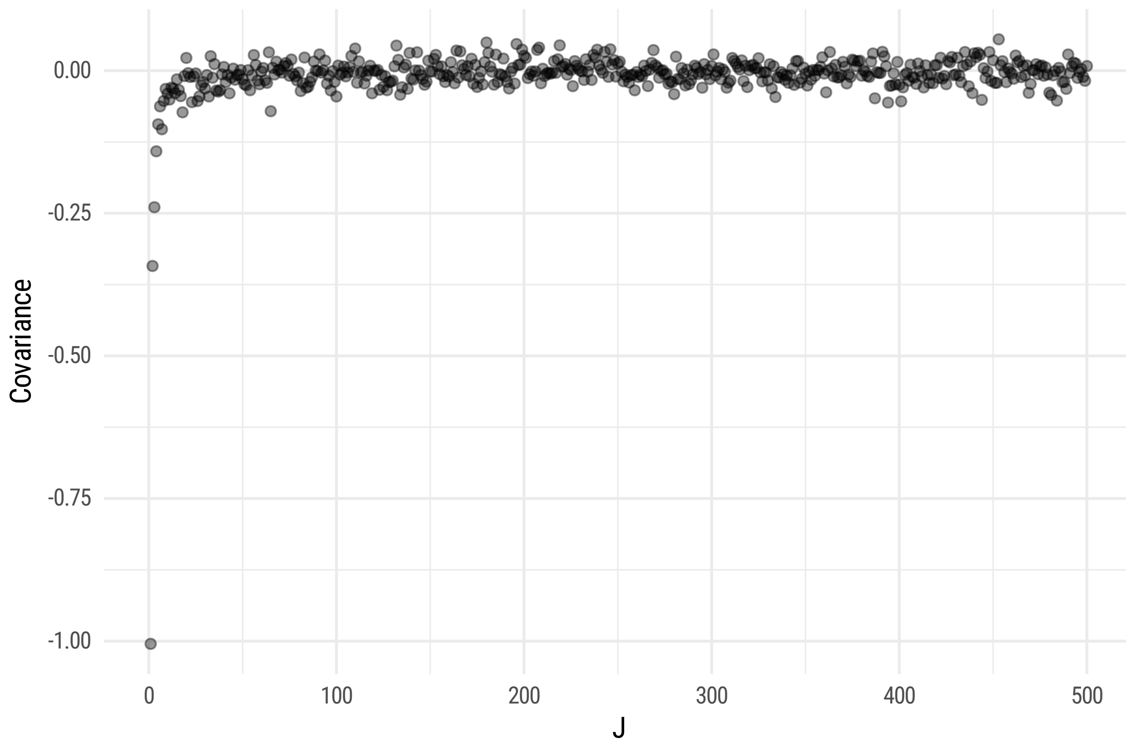

Figure 2: Behavior of the covariance as J grows larger.