Addendum to Chaganty & Joe (2006), Biometrika, pp. 197-206

Proof of Theorem 1 (iii). We will prove the result for any $d \geq 3$.

Theorem. Let $y = (y_1, \ldots, y_d)$ be a binary vector with mean $p = (p_1, \ldots, p_d) \in (0, 1)^d$, and a structured correlation matrix R. If the structure of R is AR(1) with parameter ρ, then a joint distribution for y exists if and only if

$$\max_{1 \leq i \leq (d-1)} L(p_i, p_{i+1}) \leq \rho \leq \min_{1 \leq i \leq (d-1)} U(p_i, p_{i+1}), \tag{A2}$$

where $L(a,\, b)$ and $U(a,\, b)$ are as defined on page 198 in Chaganty and Joe (2006).

Proof. The necessity is obvious. Sufficiency follows from the results in Qaqish (Biometrika, 2003, pp. 455-463). It is easy to check that if (A2) holds then λ_i given by (6) in Qaqish (2003, p. 458) will lie in the interval [0, 1] for $i = 2, \ldots, d$. Hence there exists a joint distribution for $y = (y_1, \ldots, y_d)$ in the conditional linear family.

A direct argument for the sufficiency is based on Markov chains of order 1 for binary time series discussed in Joe (1997, pp. 246-248). Let $q_i = 1 - p_i$ and $\sigma_i = (p_i\, q_i)^{1/2}$. Assume that (A2) holds. Then

$$H_i = \begin{array}{c c} \begin{array}{cc} \multicolumn{2}{c}{y_{i+1}} \\ 0 & 1 \end{array} & \\ \left[\begin{array}{cc} q_{i+1} + \rho\frac{\sigma_i \sigma_{i+1}}{q_i} & p_{i+1} - \rho\frac{\sigma_i \sigma_{i+1}}{q_i} \\ q_{i+1} - \rho\frac{\sigma_i \sigma_{i+1}}{p_i} & p_{i+1} + \rho\frac{\sigma_i \sigma_{i+1}}{p_i} \end{array}\right] & \begin{array}{c} 0 \\ 1 \end{array} \quad y_i \end{array}$$

is a transition matrix for $i = 1, 2, \ldots, (d-1)$. We can construct a joint distribution for $y = (y_1, \ldots, y_d)$ explicitly using a first order Markov chain. Let y_1 be Bernoulli with mean p_1. For $i \geq 1$, assume that the transition from y_i to y_{i+1} is governed by the transition matrix H_i. It is easy to check that the marginal of y_i is Bernoulli with mean p_i, and $\mathrm{Corr}(y_i, y_j) = \rho^{|i-j|}$ for all $1 \leq i,\, j \leq d$.

Note that, in the above, the joint distributions for $y = (y_1, \ldots, y_d)$ obtained by the conditional linear family and by the first order Markov chain, are identical.