Bayesian Analysis (2012) **7**, Number 4, pp. 801–804

Comment on Article by Kim et al.

Bruno Scarpa *

I really enjoyed reading this very interesting paper. I find human fecundity a very intriguing topic in biometrical analysis, and the authors propose a clever approach to model fecundability on the mainstream of extensions of the seminal Barrett and Marshall model. As well described in this paper, it is quite challenging to develop a good model which takes into account all the characteristics of the available data, which are very complex, so that models need to include latent and unobserved variables. I would like to congratulate the authors on their success in achieving this.

However, my role requires me to raise some points for discussion. I group them in two parts, one related to the model and the other to the real data analysis.

1 The model

I find both simple and clever the idea of using a cumulative distribution function with positive support as the link function, to relate all available explanatory variables, including acts of intercourse, to the probability of conception, as described in models (3) and (5). However, if I want to use this model, I would have difficulty in choosing the 'right' F for my data. Following Czado and Santner (1992), the authors emphasize the importance of choosing a good link function, but I did not find in the paper any suggestion or guideline aiding me, when I am defining the model (i.e., before actual data analysis), to identify the link I should use. Are there any guidelines for it?

For example, recalling the results of Czado and Santner (1992), it seems that the highest effect of the choice of the link function is observed when a skewed link function is used to fit data with actual symmetric links. So I wonder if the authors can indicate when a skewed link function is more appropriate than a symmetric one. Along these lines, it would also be interesting, in the simulation and/or application, to see how the model fits the data when a skew link function is used (considering, for example, the Box-Cox link function used by Czado and Santner 1992, or the distribution function of a skew-normal variable; see, e.g., Azzalini 1985).

A data-driven approach would further generalize the model and 'let data suggest' the right link function, – for example, by choosing a non-parametric model for it. Clearly, this might be a different model, more computationally demanding, and with different theoretical characteristics.

A second aspect in the model definition regards prior distributions. The authors choose widely dispersed priors, since they want to include very little information in addition to data. Is it possible to think about informative priors – for example, for day-specific probability parameters? In particular, considering the dataset analysed in

*Department of Statistical Sciences, University of Padua, Padua, Italy scarpa@stat.unipd.it

 DOI:10.1214/12-BA726A

the applications with LH, only 306 women were available, but there is already abundant information in the literature about the day-specific probability of conception with respect to LH (see, for example, Wilcox, Weinberg, and Baird 1995). The authors could use some of the known results from previous studies to elicit an informative prior and obtain more reliable results.

2 Applications

Many times, in the paper, the authors claim that, because the length of the fertile window is 6 days, the choice of $K = 6$ is usually assumed in the literature. I think this is not quite correct (or true). In fact, although the fertile window is 6 days, because we observe a proxy (LH or OK are two ovulation markers which include an error; see, e.g., Ecochard, Boehringer, Rabilloud, and Marret 2001), the window with positive probability of conception is longer than 6 days. In fact, K longer than 6 has been used quite frequently in the literature (see, e.g., Schwartz, MacDonald, and Heuchel 1980; Colombo and Masarotto 2000; Colombo, Mion, Passarin, and Scarpa 2006). In most fertility datsets which include cycles with a single intercourse, a small number of conceptions has been recorded when the act was in fact outside the 6-day window, requiring $K > 6$. I agree that, for computational purposes, it may be convenient to consider only the 6 days with the highest probability, but still I would not claim that those are the only ones in which conception may occur, particularly when the OK indicator (which is much less precise than LH) is used.

I also have some problems with the Bayesian model discussion. I find it curious that the final model includes explanatory variables, i.e., smoking, alchool, caffeine, and age, the 95% HDP interval of which includes 0. I would conclude that, with high probability, these variables do not affect the probability of conception, so that I would exclude them in the final model, and in particular I would not comment on the sign of their posterior mean. Conversely, because of the Bayesian approach, I would not claim parity as a 'significant' predictor: I have difficulty in understanding what 'significant' means in a Bayesian framework. In fact, this is simply a question of names and not of content, which I think is clear (but I may be biased by a frequentist background).

I think that model comparisons is another interesting aspect to be discussed. I understand that, by using all the comparison criteria (deviance, DIC and LPML), the generalized t-link (GTR) model fits the data best. I have more difficulty in agreeing that the exponential link conception model with the logit link and complementary log-log link does not fit the data well. In fact, the differences between the indicators are not so large, in view of the number of parameters involved, so we cannot be so sure about the bad fit of the Dunson and Stanford model. Note that the MCMC algorithm proposed to fit the GTR model relies on a number of Metropolis steps, while the algorithm for the Dunson and Stanford model is characterised by completely specified full conditional distributions. Therefore, this last model probably shows faster convergence and better mixing, and if the fits of the two models are really very similar, it would be preferable to choose the less computer-intensive one.

In addition, in order to better understand how the fitting of these models differs, it would be interesting to calculate some more intuitive measure of fitting. For example, since they are dealing with binary data, the authors can obtain a misclassification table or the misclassification error for the various models. They could also use more sophisticated tools, such as ROC or lift curves, to see how the models fit by changing the threshold for classification.

Lastly, a comment on the chosen ovulation markers. It is well-known that hormone-based markers are much more precise that calendar methods to identify ovulation (e.g., Ecochard, Boehringer, Rabilloud, and Marret 2001). However, it is not clear why the fact that 'the day-specific conception probability is highest on day one (-1) based on the LH method, but two days (-2) prior to estimated ovulation based on the OK method' suggests that the LH surge is a better marker for identifying the day of ovulation.

References

Azzalini, A. (1985). "A class of distributions which includes the normal ones." *Scandinavian Journal of Statistics*, 12: 171–178. 801

Colombo, B. and Masarotto, G. (2000). "Daily fecundability: first results from a new data base." *Demographic research*, 3: 5. 802

Colombo, B., Mion, A., Passarin, K., and Scarpa, B. (2006). "Cervical mucus symptom and daily fecundability: First results from a new data base." *Statistical Methods in Medical Resources*, 15: 161–180. 802

Czado, C. and Santner, T. J. (1992). "The effect of link mis-specification on binary regression inference." *Journal of Statistical Planning and Inference*, 33: 213–231. 801

Ecochard, R., Boehringer, H., Rabilloud, M., and Marret, H. (2001). "Chronological aspects of ultrasonic, hormonal, and other indirect indices of ovulation." *British Journal of Obstetrics and Gynaecology*, 108: 822–829. 802, 803

Schwartz, D., MacDonald, P. D. M., and Heuchel, V. (1980). "Fecundability, coital frequency and the viability of ova." *Population Studies*, 34: 397–400. 802

Wilcox, A. J., Weinberg, C., and Baird, D. (1995). "Timing of sexual intercourse in relation to ovulation." *New England Journal of Medicine*, 333: 1517–1521. 802