Correction on Köse T, Orman M, Ikiz F, Baksh MF, Gallagher J, Böhning D. Extending the Lincoln-Petersen estimator for multiple identifications in one source. *Statistics in Medicine* 2014;**33**: 4237-4249

T. Köse, M. Orman, F. Ikiz,* J. Gallagher,† D. Böhning‡ §

January 9, 2017

The following note refers to our paper Köse *et al.*[1]. Although the paper itself is correct, unfortunately two errors have occurred in the way the results have been implemented in the accompanying source code. In the following we point out these errors and provide the corrected code in R.

1. In section 4 of Köse *et al.* [1] a goodness-of-fit statistics was provided. Under the assumption of homogeneity, we have

$$H_0 : p_{j|0} = \frac{\lambda^j/j!}{\lambda + \lambda^2/2} \;\; for \;\; j = 1,2 \;\; and \;\; p_{j|1} = \frac{\lambda^j/j!}{1 + \lambda + \lambda^2/2} \;\; for \;\; j = 0,1,2$$

   and a test statistic

$$\chi^2 = \frac{(f_{10} - \hat{p}_{0|1}f_{1+})^2}{\hat{p}_{0|1}f_{1+}} + \frac{(f_{11} - \hat{p}_{1|1}f_{1+})^2}{\hat{p}_{1|1}f_{1+}} + \frac{(f_{12} - \hat{p}_{2|1}f_{1+})^2}{\hat{p}_{2|1}f_{1+}} + \frac{(f_{01} - \hat{p}_{1|0}f_{0+})^2}{\hat{p}_{1|0}f_{0+}} + \frac{(f_{02} - \hat{p}_{2|0}f_{0+})^2}{\hat{p}_{2|0}f_{0+}} \tag{1}$$

   with 2 df. under H_0. Unfortunately, in the R-code the goodness-of-fit statistic under the alternative has been provided. This is now corrected.

2. The second error is also an implementation error referring to part A of the online supporting information. Here, it is correctly stated that $Var(\hat{\lambda})$ can be estimated from the observed Fisher information as $-1/(\frac{\partial^2}{\partial \lambda^2} \log \ell(\hat{\lambda}, \hat{\lambda}))$, where $\log \ell(\hat{\lambda}, \hat{\lambda})$ is the log-likelihood given in (8) of Köse

*Department of Biostatistics and Medical Informatics, Faculty of Medicine, Ege University, Izmir, Turkey
†cStatistical Services Center, School of Mathematical and Physical Sciences, University of Reading, Whiteknights, Reading, RG6 6FN, U.K.
‡Southampton Statistical Sciences Research Institute, Mathematics and Medicine, University of Southampton, Southampton, SO17 1BJ, U.K.
§Correspondence to: D. Böhning, Southampton Statistical Sciences Research Institute, Mathematics and Medicine, University of Southampton, Southampton, SO17 1BJ, U.K., email:d.a.bohning@soton.ac.uk

et al.[1]. In full detail, this log-likelihood is given as

$$
\begin{aligned}
\log \ell(\lambda,\lambda) &= f_{01}\log p_{1|0} + f_{02}\log p_{2|0} + f_{10}\log p_{0|1} + f_{11}\log p_{1|1} + f_{12}\log p_{2|1} \qquad (2)\\
&= f_{01}\log\left(\frac{\lambda}{\lambda+\lambda^2/2}\right) + f_{02}\log\left(\frac{\lambda^2/2}{\lambda+\lambda^2/2}\right)\\
&\quad + f_{10}\log\left(\frac{1}{1+\lambda+\lambda^2/2}\right) + f_{11}\log\left(\frac{\lambda}{1+\lambda+\lambda^2/2}\right) + f_{12}\log\left(\frac{\lambda^2/2}{1+\lambda+\lambda^2/2}\right)\\
&= f_{+1}\log(\lambda) - f_{0+}\log(\lambda+\lambda^2/2) + 2f_{+2}\log(\lambda) - f_{1+}\log(1+\lambda+\lambda^2/2) \qquad (3)
\end{aligned}
$$

where $f_{+1} = f_{01} + f_{11}$, $f_{+2} = f_{02} + f_{12}$, $f_{0+} = f_{01} + f_{02}$ *and* $f_{1+} = f_{10} + f_{11} + f_{02}$.

It follows that

$$
\begin{aligned}
\frac{\partial}{\partial\lambda}\log\ell(\lambda,\lambda) &= \frac{f_{+1}}{\lambda} - f_{0+}\frac{(1+\lambda)}{(\lambda+\lambda^2/2)} + \frac{2f_{+2}}{\lambda} - f_{1+}\frac{(1+\lambda)}{(1+\lambda+\lambda^2/2)}\\
\frac{\partial^2}{\partial\lambda^2}\log\ell(\lambda,\lambda) &= -\frac{f_{+1}}{\lambda^2} - \frac{2f_{+2}}{\lambda^2} - f_{0+}\left(\frac{(\lambda+\lambda^2/2)-(1+\lambda)^2}{(\lambda+\lambda^2/2)^2}\right) - f_{1+}\left(\frac{(1+\lambda+\lambda^2/2)-(1+\lambda)^2}{(1+\lambda+\lambda^2/2)^2}\right). \qquad (4)
\end{aligned}
$$

Instead of (4) the following was implemented:

$$
\begin{aligned}
&- f_{+0}\left(\frac{(1+\lambda+\lambda^2/2)-(1+\lambda)^2}{(1+\lambda+\lambda^2/2)^2}\right) - f_{+1}\left(\frac{1}{\lambda^2} + \frac{(1+\lambda+\lambda^2/2)-(1+\lambda)^2}{(1+\lambda+\lambda^2/2)^2}\right) \qquad (5)\\
&- f_{+2}\left(\frac{2}{\lambda^2} + \frac{(1+\lambda+\lambda^2/2)-(1+\lambda)^2}{(1+\lambda+\lambda^2/2)^2}\right),
\end{aligned}
$$

where $f_{+0} = f_{00} + f_{10}$, $f_{+1} = f_{01} + f_{11}$ and $f_{+2} = f_{02} + f_{12}$.

This error has now been corrected as well.

The R-code with both errors corrected is now available in the online material. The equivalent web-software `http://biostat.vecdiaytac.com/LPMultiple.aspx` has been updated as well.

Acknowledgment

We are most grateful to Dr Rattana Lerdsuwansri (Thammasat University Bangkok) for pointing out these errors to us.

Reference

1. Köse T, Orman M, Ikiz F, Baksh MF, Gallagher J, Böhning D. Extending the Lincoln-Petersen estimator for multiple identifications in one source. *Statistics in Medicine* 2014;**33**: 4237-4249, DOI: 10.1002/sim.6208.