Sampling error variance in weighted three-day average actiwatch measures of sleep duration

Technical appendix to "Self-reported and measured sleep duration: how similar are they? The CARDIA Sleep Study" by DS Lauderdale, KL Knutson, LL Yan, K Liu, and PJ Rathouz, February, 2008.

Introduction

As described in the paper, as a preliminary analysis, we treated the three consecutive days of actiwatch sleep duration data as a stratified simple random sample of the past thirty days with equal day-to-day variance for weekdays and weekends. Two potentially serious concerns with this approach are the following: (i) sleep duration measures on these three days, because they occur consecutively, might be autocorrelated, and (ii) the daily-variance in sleep duration might be different on the weekends than on the weekdays. Positive autocorrelation and larger variance on the weekends will mean that our initial error variance estimates (in using the 3-day weighted average as an estimate of the true 30-day weighted average) are biased low and the corresponding reliability estimates (for the 3-day relative to the averages) are biased high. In the measurement error regression models relating subjective to objective measures, this would then mean that slopes, correlations and variance explained between the subjective and objective measures are **under-corrected** for measurement error in the objective measure, and hence **biased low** as well.

Negative autocorrelation is also possible; e.g., one might compensate for little sleep on one night by increased sleep in one or two of the following nights. Negative autocorrelation would mean that our reliability estimates in the paper are biased low and hence that the strength of association between subjective and objective measures is biased high. Considering that the "conservative estimates" in our paper are those that produce the highest correlation between subjective and objective measures and the lowest bias in the subjective measures relative to the objective ones, we will concentrate here primarily on the problem of positive autocorrelation and increased weekend variance.

Fortunately, each participant supplied two series of three-day sequences of sleep measures, one year apart. This second year of data permits us to get a handle on the degree to which these two problems could be affecting our results, and to estimate correction factors for the estimated error variance from our initial estimates due to autocorrelation and different weekend than weekday variance.

Procedure

We fitted a mixed effects linear regression model (Laird and Ware, 1992) of the following form to both years' data:

$$Y_{ijt} = \beta_1 I(t = 1, 2) + \beta_2 I(t = 3) + U_{i1} I(t = 1, 2) + U_{i2} I(t = 3) + \epsilon_{ijt} \ . \qquad (1)$$

Here i indexes subject, j indexes year ($j = 1, 2$), t indexes day of measurement ($t = 1, 2, 3$), weekdays nights are $t = 1, 2$, the weekend night is $t = 3$, $I(\cdot)$ is an indicator variable, and Y_{ijt} is the actiwatch sleep duration measure for subject i, year j and day t. The fixed parameters β_1 and β_2 represent the population average sleep duration for weekday and weekend nights across all subjects and both years. The random effects $U_i = (U_{i1}, U_{i2})'$ represent subjectspecific deviations from this average, again averaged over the two years. The long-term average weekday sleep duration for subject i is therefore given by $\beta_1 + U_{i1}$, while that for weekends is $\beta_2 + U_{i2}$. We assume that U_i is bivariate normally distributed across subjects with mean zero and variance-covariance matrix $\mathrm{cov}(U_i)$.

We also assume that, given U_i, the two years' data on a given subject are uncorrelated, i.e., $\mathrm{cov}(\epsilon_{ijt}, \epsilon_{ij't'}) = 0$ for $j \neq j'$. This is reasonable, as the role of U_i in the model is precisely to capture long-range correlation among observations on the same subject. Then

$$\Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix},$$

where $\Sigma_1 = \mathrm{cov}(\epsilon_{i1})$, $\Sigma_2 = \mathrm{cov}(\epsilon_{i2})$, and $\epsilon_{ij} = (\epsilon_{ij1}, \epsilon_{ij2}, \epsilon_{ij3})'$, $j = 1, 2$. This model permits identification and estimation of $\mathrm{cov}(U_i)$ and of the variance-covariance matrix $\Sigma = \mathrm{cov}(\epsilon_i)$, where $\epsilon_i = (\epsilon'_{i1}, \epsilon'_{i2})'$. As the data used in the 30-day subjective-objective analyses were from the first year of the study, a key quantity of interest is Σ_1.

Fitting model (1) to the entire data set (not stratified by any demographic variables), we obtained

$$\widehat{\Sigma}_1 = \begin{pmatrix} 1.3361 & 0.1294 & 0.1794 \\ 0.1294 & 1.3258 & 0.02427 \\ 0.1794 & 0.02427 & 2.2631 \end{pmatrix},$$

or, in correlation form,

$$\begin{pmatrix} 1.0000 & 0.09725 & 0.1031 \\ 0.09725 & 1.0000 & 0.01401 \\ 0.1031 & 0.01401 & 1.0000 \end{pmatrix}.$$

Note importantly that the correlation from day to day is relatively small, that the variance across the two weekdays is almost constant, and that the variance on the weekend is considerably larger than that on the weekdays. To proceed, we averaged the nearly-equal variances on the two weekdays to obtain

$$\widehat{\Sigma}_1 = \begin{pmatrix} 1.3310 & 0.1294 & 0.1794 \\ 0.1294 & 1.3310 & 0.02427 \\ 0.1794 & 0.02427 & 2.2631 \end{pmatrix}.$$

Suppose as a working model, we write

$$\epsilon_{ijt} = V_{ij1} I(t = 1, 2) + V_{ij2} I(t = 3) + Z_{ijt} \tag{2}$$

where the random effects $V_{ij} = (V_{ij1}, V_{ij2})'$ represent month-specific deviations from subject i's average sleep duration on weekdays and weekends, and the Z_{ijt}'s are daily deviations from that monthly average. Let $G = \text{cov}(V_{ij})$ and $R_1 = \text{cov}(Z_{i1})$ where $Z_{i1} = (Z_{i11}, Z_{i12}, Z_{i13})'$.

The variances and covariances R_1 of the Z_{ijt}'s are the relevant quantities in determining the error variance of the three-day weighted averages as estimates of the thirty day weighted averages for each subject. This is because, under model (1)-(2), the true thirty day weighted average (for subject i, year 1) is

$$w_1(Y_{i11} - Z_{i11}) + w_2(Y_{i12} - Z_{i12}) + w_3(Y_{i13} - Z_{i13}) \ ,$$

whereas what is estimated by the 3-day average is

$$w_1 Y_{i11} + w_2 Y_{i12} + w_3 Y_{i13} \ .$$

The error variance of this estimate is

$$\mathrm{var}(w_1 Z_{i11} + w_2 Z_{i12} + w_3 Z_{i13}) = w' R_1 w \ ,$$

where $w = (w_1, w_2, w_3)'$. The weights are $w_1 = w_2 = 5/14$ and $w_3 = 4/14$, corresponding to our three-day weighted average being a stratified random sample of two weekdays and one weekend night. Note that under model (1)-(2), the total variance of weighted averages $(w_1 Y_{i11} + w_2 Y_{i12} + w_3 Y_{i13})$ is $w'\Sigma_1 w$. Under our fitted model, this is estimated as $w'\widehat{\Sigma}_1 w = 1.355$ h^2, which compares very well with the simple sample variance of the weighted average sleep duration (1.341 h^2).

Of course, R_1 is not identifiable, but with the constraints that both G and R_1 must be positive definite covariance matrices, we can learn something about the range of possible values of $w' R_1 w$ given estimate $\widehat{\Sigma}_1$. Specifically, note that

$$R_1 = \Sigma_1 - XGX', \tag{3}$$

where

$$X = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(The rows of X correspond to days $t = 1, 2, 3$, the two columns being indicator functions for weekday or weekend nights.)

Range of possible values for estimated error variance

Using (3), we maximized error variance $w'R_1w$ over all possible values of G holding $\widehat{\Sigma}_1$ constant. The maximum occurred at $G = 0$, in which case $R_i = \widehat{\Sigma}_1$. This corresponds to the situation where there is no within-subject variability in sleep duration on the scale of month-to-month or longer. Rather, variability for each subject is more on the scale of week-to-week and hence shows up as positive autocorrelation between observations made on consecutive days. Also, in this situation, the increased variability on the weekend nights is not due to month-to-month variability in weekend sleep duration, but rather operates on a smaller time scale. This situation represents the worst possible scenario with respect to error variance due to concerns (i) and (ii) outlined in the first paragraph. The estimated error variance in this situation is 0.599 h^2, as compared to an initial estimate of 0.402 h^2 assuming no auto-correlation and equal weekend and weekday variance, representing an underestimate of about 33% in error variance.

It is also possible that the month-to-month variability in V_{ij} is such that the Z_{ijt}'s are uncorrelated or even negatively correlated and that the increased variability on the weekend nights is accounted for by month to month variability in V_{ij}. It turns out to be more technically difficult to minimize $w'R_1w$ over all G than it is to maximize it, but we can select a few key values of G to given a sense of how things might work out. In the first of these, we choose G so that the the daily variance is constant across all three days and the first and second and first and third days are uncorrelated. This yields

$$R_1 = \begin{pmatrix} 1.2016 & 0.0000 & 0.0000 \\ 0.0000 & 1.2016 & -0.1551 \\ 0.0000 & -0.1551 & 1.2016 \end{pmatrix}$$

with an error variance of $w'R_1w = 0.373$ h^2. Note that this yields a small negative auto-correlation between nights two and three. In the second one, we choose G to render the daily variance even smaller, and allow the correlation between nights one and two to also be negative. This yields

$$R_1 = \begin{pmatrix} 1.1516 & -0.0500 & 0.0000 \\ -0.0500 & 1.1516 & -0.1551 \\ 0.0000 & -0.1551 & 1.1516 \end{pmatrix}$$

with an error variance of $w'R_1w = 0.343$ h^2. In this situation, our initial estimates would be biased high by about 17%.

Estimated error variance under inflated weekend variance

We now turn to the more restricted setting wherein there is within-subject variability in sleep duration on the month-to-month scale which is **similar** for weekdays and weekends, and so does not account for the observed increased weekend variance in $\widehat{\Sigma}_1$. This increased weekend variance therefore must figure into the error variance of the 3-day average.

We first consider the setting where the month-to-month variability accounts for the autocorrelation in $\widehat{\Sigma}_1$, so that there is no autocorrelation present in the 3-day average as an estimate of the 30-day average. This yields

$$R_1 = \begin{pmatrix} 1.202 & 0.000 & 0.078 \\ 0.000 & 1.202 & -0.077 \\ 0.078 & -0.077 & 2.134 \end{pmatrix}$$

(the covariances between nights 1 and 3 and nights 2 and 3 cancel in the error variance calculation). The error variance in this situation is estimated as $w'R_1w = 0.481$ h^2, which would mean that our initial estimates under no autocorrelation and equal weekday and weekend variance would be biased low by about 16%.

Now suppose that positive autocorrelation exists in the 3-day average as an estimate of the 30-day average. The most extreme case for this in terms of error variance is simply that where $G = 0$ and $R_1 = \widehat{\Sigma}_1$. We have already seen that the estimated error variance in this situation is 0.599 h^2, 25% greater than the case of no autocorrelation.

Finally, consider the case of negative autocorrelation in the 3-day average as an estimate of the 30-day average. How strongly can this effect the error variance? Based only on the estimated value of Σ_1, G could be chosen such that the resulting R_1 contains rather large negative autocorrelation. A reasonable choice specifies G as accounting for all positive autocorrelation, in particular, the largest autocorrelation in the observed $\widehat{\Sigma}$, which is that between day 1 and day 3. This yields

$$R_1 = \begin{pmatrix} 1.1516 & -0.05000 & 0.00000 \\ -0.0500 & 1.15155 & -0.15513 \\ 0.0000 & -0.15513 & 2.08370 \end{pmatrix}.$$

The error variance in this setting is 0.419 h^2, 13% less than the case of no autocorrelation.

Sensitivity analyses

We first ran our measurement error models using our nominal initial error variance estimate of 0.402 h^2. As sensitivity analyses for how the unknown error variance could be affecting our results, we reran these models inflating our initial estimate (i) by a factor of $(0.599/0.402) = 1.49$ to account for a potential underestimate due to ignoring positive autocorrelation and increased weekend variance, (ii) by a factor of $(0.481/0.402) = 1.20$ to account for a potential underestimate due to increased weekend variance only, (iii) by a factor of $(0.419/0.402) = 1.04$ to account for a potential underestimate due to ignoring increased weekend variance, even assuming negative autocorrelation. We also (iv) deflated the error variance by a factor of $(0.343/0.402) = 0.85$ to account for a potential overestimate due to negative autocorrelation. The results are compared in the following table. (numbers in parentheses are 95% bootstrap BCa confidence intervals, 2000 replications). The fourth line is what we ultimately reported as the first entry in Table 2 of the paper; the third and fifth lines are the sensitivity analyses reported in the paper.

Model	Err. Var. Inflation	Reliab.	Self-reported vs. Measured Sleep	Slope	Correlation
NoWV, NegAC	0.85	0.74	0.80 (0.72, 0.88)	0.48 (0.35, 0.61)	0.43 (0.34,0.52)
NoWV, NoAC*	1.0	0.70	0.80 (0.72, 0.88)	0.51 (0.37, 0.66)	0.45 (0.35, 0.54)
WV, NegAC***	1.04	0.69	0.80 (0.71, 0.88)	0.52 (0.37, 0.67)	0.45 (0.35, 0.55)
WV, NoAC**	1.20	0.64	0.80 (0.71, 0.88)	0.56 (0.39, 0.73)	0.47 (0.36, 0.57)
WV, PosAC***	1.49	0.55	0.79 (0.70, 0.88)	0.65 (0.42, 0.88)	0.50 (0.38, 0.62)

(No)WV=(No) excess weekend variability.

Neg/No/PosAC=Negative/No/Positive auto-correlation.

*These results are obtained under our initial analysis.

**These results are what is reported in Table 2 of the paper.

***These results are reported in the paper as sensitivity analyses.

Conclusion

Because one could make an argument that the auto-correlation from night-to-night is as likely to be negative as it is to be positive, we have chosen to maintain the assumption of no autocorrelation for our analyses for the revised paper. We do believe, however, that the additional weekend variance observed in $\widehat{\Sigma}_1$ should be accounted for in our analyses. In the final version of the paper, therefore, we have inflated the nominal estimate of the error variance in the 3-day average relative to the 30-day average by a factor of $(0.481/0.402) = 1.20$ for our main analyses and for all stratified analyses presented in Table 2. (Note: We did not re-estimate the variance-covariance matrix $\widehat{\Sigma}_1$ for each stratum, as this would have been computationally quite cumbersome.) For the primary, unstratified analysis in the paper, we also report as a sensitivity analyses, the results under negative and positive autocorrelation (error variance inflation by 1.04 and 1.49), again assuming that the additional weekend variance in $\widehat{\Sigma}_1$ should figure into the error variance calculation.