A Simulation Study to Compare Two-Stage Models and Joint Models

Samantha Bothwell | BIOS 7721 : Joint Modeling | March 5th, 2021

Introduction

Joint and two-stage models are are two methods of understanding the association between longitudinal outcomes and survival outcomes.

Methods

Simulation All analysis was performed in R, version 4.0.2. The simulation was conducted over 500 data sets, with 250 individuals each. Longitudinal outcomes were simulated under the following linear mixed model framework :

$$y_{ij} = -0.2 + 0.3t + b_{0i} + b_{1i}t + \epsilon_{ij} = m_i(t) + \epsilon_{ij}, \epsilon_{ij} \sim N(0, (0.5)^2), \quad (b_{0i}, b_{1i})' \sim N(0, A) \text{ with } A = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.3 \end{bmatrix}$$

Measurement times were simulated randomly from a poisson process with an anverage of one measurement per year. Individuals had a maximum of 10 measurements, including baseline.

Survival times were simulated under the following cox model framework with a constant baseline hazard ($h_0(t) = 0.1$) :

$$h_i(t) = h_0(t)exp\{\alpha m_i(t)\} = 0.1exp\{0.6 \cdot (-0.2 + 0.3t + b_{0i} + b_{1i}t)\}$$

Censoring times were simulated randomly from a uniform(0, 10) distribution. If the censoring time occured before the survival time, the individual was classified as censored at that time, otherwise the individual was classified as having experienced death at the survival time. Any measurements after the censored or survival time, whichever came first, were removed.

The linear mixed model, for both the joint and two-stage processes, was fit with the lme function in the nlme R package, version 3.1-148. The cox model was fit with the coxph function in the survival R package, version 3.1-12. The joint model was fit assuming a propotional hazards model with a Weibull baseline hazard. The pseudo-adaptive Gauss-Hermite rule was used to improve computation time. The model was fit using the JM R package, version 1.4-8.

To fit the two-stage model, we created start and stop times for intervals between each visit, which indicated the interval during with the event of death occurred. The Cox model was then fit using the subject-specific predictions of the true, unobserved longitudinal outcome values.

Comparison metrics We compared the accuracy of the two modeling methodologies with the following variables:

- *Empirical Bias:* The estimated difference between the average parameter value across all 500 simulations and the true parameter value.
- *Asymptotic Standard Error:* The average model standard errors for the association, α, and β parameters.
- *Empirical Standard Error:* The esimated standard deviation of each of the parameter value across all 500 simulations.
- *Mean Square Error:* The average sum of squared differences between each estimated parameter and the true parameter value across all 500 simulations.
- *95% CI Coverage Rates:* The proportion of 95% confidence intervals, for the association, α, and β parameters, that contain the true parameter value.

In additon, the total model run time for each modeling method was computed and their average time was compared.

Results

Table 1, below, shows the comparison metric values by model. The magnitude of bias is generally smaller for the joint model compared to the two stage model.

Table 1. Model Accuracy Metrics

	2 Stage Model						Joint Model					
	Assoc	B0	B1	Sigma	A11	A22	Assoc	B0	B1	Sigma	A11	A22
Bias	-0.0091	0.0069	-0.0279	0.00090	0.00021	-0.0066	0.0081	-0.0015	0.00212	0.00016	-0.0018	-0.0013
Asy SE	0.1072	0.0400	0.0278	NA	NA	NA	0.1014	0.0399	0.02882	NA	NA	NA
Emp SE	0.1013	0.0393	0.0285	0.01486	0.03622	0.0244	0.0977	0.0393	0.02975	0.01481	0.0364	0.0247
MSE	0.0103	0.0016	0.0016	0.00022	0.00131	0.0867	0.0096	0.0015	0.00089	0.00022	0.0013	0.0898
95% CI Coverage	0.9580	0.9500	0.8260	NA	NA	NA	0.9700	0.9600	0.94400	NA	NA	NA

The average run time for the 2 stage model was 0.17 seconds compared to 7.32 seconds for the joint model.

Discussion

References