Test how does the model fit

2020-03-08

For a linear mixed effect model

$$Y = X\beta + Zb + \epsilon \tag{1}$$

- Y is the outcome vector, with dimension $n \times 1$
- X, Z are design matrices, with dimension $n \times p$
- β is the fixed effect, with dimension $p \times 1$
- b is the random effect, $b \sim N(0, D)$, D is a $p \times p$ symmetric matrix
- $\epsilon \sim N(0, \sigma^2 I)$, I is $n \times n$, *epsilon* $\perp\!\!\!\perp D$

The **covariance matrix** of Y is

$$cov(Y) = cov(Zb) + \sigma^2 I = ZDZ' + \sigma^2 I$$

We may also rewrite the model as

$$Y = X\beta + \epsilon^\star, \epsilon^\star = Zb + \epsilon \sim N(0, V), V = ZDZ' + \sigma^2 I \tag{2}$$

The matrix V is symmetric, and $V = V^{\frac{1}{2}}(V^{\frac{1}{2}})'$, therefore,

$$V^{-\frac{1}{2}}Y = V^{-\frac{1}{2}}X\beta + V^{-\frac{1}{2}}\epsilon^\star \tag{3}$$

$$cov(V^{-\frac{1}{2}}\epsilon^\star) = V^{-\frac{1}{2}}V(V^{-\frac{1}{2}})' = V^{-\frac{1}{2}}V^{\frac{1}{2}}(V^{\frac{1}{2}})'(V^{-\frac{1}{2}})' = I_{n,n}$$

Recall that for a linear regression,

$$Y = X\beta + \epsilon, \epsilon \sim N(0, I)$$

the estimation of β has a formular

$$\hat{\beta} = (X'X)^{-1}X'Y$$

Therefore, the **estimation of** β from equation (3) can be written as

$$\hat{\beta} = ((V^{-\frac{1}{2}}X)'V^{-\frac{1}{2}}X)^{-1}(V^{-\frac{1}{2}}X)'V^{-\frac{1}{2}}Y = (X'V^{-1}X)^{-1}X'V^{-1}Y$$

The **covariance of** $\hat{\beta}$ is

$$
\begin{aligned}
cov(\hat{\beta}) =& cov((X'V^{-1}X)^{-1}X'V^{-1}Y) \\
=& (X'V^{-1}X)^{-1}X'V^{-1}cov(Y)[(X'V^{-1}X)^{-1}X'V^{-1}]' \\
=& (X'V^{-1}X)^{-1}X'V^{-1}V[(X'V^{-1}X)^{-1}X'V^{-1}]' \\
=& (X'V^{-1}X)^{-1}X'V^{-1}X(X'V^{-1}X)^{-1} \\
=& (X'V^{-1}X)^{-1}, V = ZDZ' + \sigma^2 I
\end{aligned}
$$

Back to our linear mixed effect model (since the drug goup and placebo group have similar parameter settings, we may just consider one group as an example and the following coefficients do not have group labels)

$$Y_i = X_i(\beta + b_i + \Gamma w_i) + \epsilon_i$$

The fixed effect is $X_i(\beta + \Gamma w_i)$, the random effect is $X_i b_i$, the random error is ϵ_i. It may be hard to test β and Γ separately. Therefore, let's test the total fixed effect $X_i(\beta + \Gamma w_i)$ first.

Suppose n is the number of subjects in each treatment group (drg, pbo). β, Γ have dimensions 3×1. The w has dimension $n \times 1$ Then let's rewrite $Y_{7\times1}, X_{7\times3}, \beta, \Gamma$ as

$$Y^\star = \begin{pmatrix} Y_1 \\ Y_2 \\ ... \\ Y_n \end{pmatrix}_{7n\times1} , X^\star = \begin{pmatrix} X_1 & 0 & ... & 0 \\ 0 & X_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & 0 & X_n \end{pmatrix}_{7n\times3n}$$

$$\beta^\star = \begin{pmatrix} \beta \\ \beta \\ ... \\ \beta \end{pmatrix}_{3n\times1} , \Gamma^\star = \begin{pmatrix} \Gamma & 0 & ... & 0 \\ 0 & \Gamma & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & 0 & \Gamma \end{pmatrix}_{3n\times n}$$

$$b^\star = \begin{pmatrix} b_1 \\ b_2 \\ ... \\ b_n \end{pmatrix}_{3n\times1} \sim N(0, D^\star), D^\star = \begin{pmatrix} D & 0 & ... & 0 \\ 0 & D & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & 0 & D \end{pmatrix}_{3n\times3n}$$

The model is

$$Y^\star_{7n\times1} = X^\star_{7n\times3n}(\beta^\star_{3n\times1} + \Gamma^\star_{3n\times n}w_{n\times1}) + X_{7n\times3n}b^\star_{3n\times1} + \epsilon^\star_{7n\times1}$$

And

$$\hat{\beta}^\star + \hat{\Gamma}^\star w \sim MVN(\beta^\star + \Gamma^\star w, V^\star_{\text{fixed}})$$

To test the estimation of fixed effect, the test statistic is calculated by

$$\mathcal{X}^2 = (\hat{\beta}^\star + \hat{\Gamma}^\star w - \beta^\star + \Gamma^\star w)' V^{\star-1}_{\text{fixed}} (\hat{\beta}^\star + \hat{\Gamma}^\star w - \beta^\star + \Gamma^\star w)$$

$$V^\star_{\text{fixed}} = (X^{\star\prime}(X^\star D^\star X^{\star\prime} + \sigma I)^{-1} X^\star)^{-1}$$

However, with larger σ, the test statistic $\mathcal{X}^2$ may not increase dramatically or may not increase. This is because, with a larger σ, the covariance of β gets larger, however, its inverse gets smaller; the estimation of fixed effect may have bigger bias, i.e. $|\hat{\beta}^\star + \hat{\Gamma}^\star w - \beta^\star + \Gamma^\star w|$ increases.

Therefore, within the formular of $\mathcal{X}^2$, we may think that the values of the numerator and denominator increase at the same time. It is hard to say whether the ratio gets larger or not. This may be one of the reason that we did not observe much differences when we increased the random error σ.

To show that, we run simulations with dimension of baseline covariances equal to 2, 4, 8, 16; angles between Γ_1 and Γ_2 equal to $0, \frac{\pi}{3}, \frac{2\pi}{3}, \pi$; variance of the random error σ^2 equal to 1, 4, or 100.

The fixed effect βs are chosen as (to avoid the intercept is too big to dominate the vector, I set them to be close to each other)

$$\beta_{drg} = (1, -2.3, 1.5)', \beta_{pbo} = (1.1, -1.9, 1)'$$

The covariances of random effect b are set as

$$D_{drg} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1 \end{pmatrix}, D_{pbo} = \begin{pmatrix} 4 & 1.73 & 1 \\ 1.73 & 1.75 & 0.43 \\ 1 & 0.43 & 1.25 \end{pmatrix}$$

A dataset is generated and then $\hat{\alpha}$, $\hat{w}$, $\hat{\beta}, \hat{\Gamma}$ are estimated. The chisquare test statistic is then calculated.

The procedures run with 100 repetitions to reduce random bias.

Results

The table of test statistic

$$\mathcal{X}^2 = (\hat{\beta}^\star + \hat{\Gamma}^\star w - \beta^\star + \Gamma^\star w)' V^{\star-1}_{\text{fixed}} (\hat{\beta}^\star + \hat{\Gamma}^\star w - \beta^\star + \Gamma^\star w)$$

Table 1: Table of test statistic

| p | sigma | angle | Drug chisquare | | Drug | Placebo chi | | Placebo | |
			mean	sd	det-ratio	mean	sd	det-ratio	cosine
2	1	0	6.48	3.09	0.79	8.51	3.09	0.72	0.96
		60	6.63	3.20	0.79	11.43	3.20	0.73	0.95
		120	6.10	2.96	0.79	7.82	2.96	0.72	0.98
		180	6.12	2.94	0.79	7.45	2.94	0.72	0.98
	2	0	6.54	3.33	0.75	8.39	3.33	0.45	0.94
		60	6.37	3.26	0.74	11.02	3.26	0.47	0.95
		120	6.02	2.97	0.74	8.33	2.97	0.46	0.98
		180	6.09	2.99	0.74	7.31	2.99	0.46	0.97
	10	0	5.79	3.51	0.18	6.46	3.51	0.07	0.97
		60	5.67	3.40	0.17	7.60	3.40	0.08	0.98
		120	5.60	3.40	0.17	6.76	3.40	0.07	0.99
		180	5.89	3.64	0.18	6.65	3.64	0.09	0.96
4	1	0	7.55	4.27	0.83	10.90	4.27	0.72	0.92
		60	8.48	4.95	0.83	21.48	4.95	0.73	0.87
		120	7.07	4.31	0.82	10.35	4.31	0.72	0.96
		180	7.14	4.33	0.82	8.93	4.33	0.72	0.95
	2	0	8.03	4.30	0.78	9.88	4.30	0.46	0.89
		60	9.19	6.17	0.77	24.60	6.17	0.48	0.85
		120	7.31	4.26	0.77	11.82	4.26	0.47	0.94
		180	7.39	4.31	0.77	8.62	4.31	0.46	0.93
	10	0	6.61	4.19	0.16	6.65	4.19	0.08	0.98
		60	6.40	4.08	0.19	7.03	4.08	0.08	0.99
		120	6.54	4.06	0.17	8.25	4.06	0.07	0.97
		180	6.52	4.13	0.19	6.61	4.13	0.07	0.98
8	1	0	8.79	4.45	0.82	19.57	4.45	0.69	0.80
		60	10.99	7.94	0.80	43.48	7.94	0.73	0.75
		120	7.11	3.49	0.81	16.67	3.49	0.69	0.89
		180	7.27	3.43	0.81	13.36	3.43	0.68	0.88
	2	0	9.23	5.53	0.78	15.23	5.53	0.44	0.78
		60	9.63	6.38	0.77	33.85	6.38	0.52	0.80
		120	7.20	3.62	0.77	17.72	3.62	0.49	0.88
		180	7.67	3.72	0.77	11.20	3.72	0.45	0.85
	10	0	5.74	3.54	0.18	6.00	3.54	0.11	1.00
		60	5.77	3.54	0.21	6.41	3.54	0.09	0.99
		120	5.74	3.53	0.20	6.23	3.53	0.10	1.00
		180	5.73	3.53	0.18	5.99	3.53	0.09	1.00
16	1	0	11.33	4.60	0.80	24.92	4.60	0.70	0.74
		60	12.46	4.99	0.79	47.32	4.99	0.74	0.71
		120	8.48	3.88	0.78	20.87	3.88	0.68	0.85
		180	8.84	3.89	0.78	17.23	3.89	0.66	0.83
	2	0	9.91	5.27	0.73	15.13	5.27	0.41	0.80
		60	9.53	6.26	0.72	29.19	6.26	0.42	0.84
		120	7.79	3.85	0.72	17.70	3.85	0.44	0.89
		180	8.50	4.15	0.71	12.20	4.15	0.38	0.85

10	0	5.69	3.16	0.16	5.61	3.16	0.10	1.00
	60	5.69	3.16	0.21	5.62	3.16	0.08	1.00
	120	5.69	3.16	0.18	5.62	3.16	0.07	1.00
	180	5.68	3.16	0.21	5.61	3.16	0.07	1.00

where the mean and standard deviation of the chisquare test statistics are presented in the above table. How did the random effect fit is evaluated as the raito of determinates of true D and estiamted $\hat{D}$ matrix.

$$\text{det-ratio} = \frac{min(|D|, |\hat{D}|)}{max(|D|, |\hat{D}|)}$$

It seems that the estimation of the fixed effects is still well even when we increase the random error a lot.

Prediction

Change score approach

To make a decision function, the straightforward idea is fitting a linear regression model, with the outcome difference during the follow-up as the outcome and biosignatures as the independent variables, that is, ignore the longtiudnatla sturcture and use the last - first score as the outcome and the baseline biosignarues serve as independent varables. That is

$$\Delta Y_{i,drg} = Y_{i,7,drg} - Y_{i,0,drg}, \Delta Y_{i,pbo} = Y_{i,7,pbo} - Y_{i,0,pbo}$$

where $\Delta Y_{i,drg}$ is the outcome differences between the first measure and last measure of the ith subject under intervention, while $\Delta Y_{i,pbo}$ is the ith subject's counterfactor outcome differences if he or she takes placebo. If $\Delta Y_{i,drg} > \Delta Y_{i,pbo}$, the patient has a larger decrease in the HRSD when treated with drug, which means that the drug has a better effect than the placebo for that patient and he or she should be assigned into the treatment group. On the other hand, if $\Delta Y_{i,drg} < \Delta Y_{i,pbo}$, the patient has less recovery when treated with drug than had placebo, therefore, the patient does not need drug treatment, since it will not improve the patient's outcome than placebo. By fitting the linear regression

$$\Delta Y_{i,drg} = \gamma_{drg} X, \Delta Y_{i,pbo} = \gamma_{drg} X$$

we can get the estimations of the coefficients $\hat{\gamma}_{drg}, \hat{\gamma}_{drg}$, and the associated estimated outcome difference can be evaluated as:

$$\Delta \hat{Y}_{i,drg} = \hat{\gamma}_{drg} X, \Delta \hat{Y}_{i,pbo} = \hat{\gamma}_{drg} X$$

To estimate the prediction effect of this decision rule, we would like to check whether the esitmated changes of outcome for a subject is consistent with the true changes, which is:

Table 2: Concordance of Prediction

	$\Delta Y_{i,drg} > \Delta Y_{i,pbo}$	$\Delta Y_{i,drg} < \Delta Y_{i,pbo}$
$\hat{\Delta} Y_{i,drg} > \hat{\Delta} Y_{i,pbo}$	1	0
$\hat{\Delta} Y_{i,drg} < \hat{\Delta} Y_{i,pbo}$	0	1

which we defined it as "concordance of group assignment". The percentage of the concordance is used to measure the performance.

Longitudinal decision rule and average tangent slope

With our previous method, the outcome is a fitted with a quadratic trajectory and is a vector instead of a scaler. Therefore, we need a measures to extract the features of the quadratic curves. The approach, average tangent slope, is considered. We use the average tangent slope as the measure of a trajectory, where

$$ATS = \frac{1}{b-a} = \int_a^b y'(t)dt = \frac{y(b) - y(a)}{b-a}$$

which is the difference between the last time score and the first time score, over that length of the time period. The associated concordance can be calculated as:

Table 3: Concordance of Prediction

	$ATS_{drg} > ATS_{pbo}$	$ATS_{drg} < ATS_{pbo}$
$\hat{ATS}_{drg} > \hat{S}_{pbo}$	1	0
$\hat{ATS}_{drg} > \hat{S}_{pbo}$	0	1

The following plots shows the cosine similarity of $\hat{\alpha}$ and α_0 across different parameter settings.
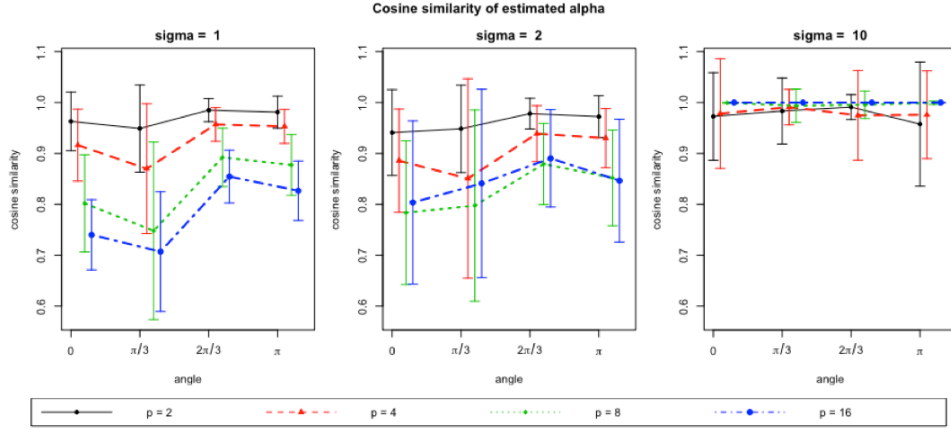
The concordances of the change score method and the longitudinal decision rule method are presenting below:

The comparison of agreement of intervention group assignement

To make it more clear, the tables to draw the above plots are showing below:
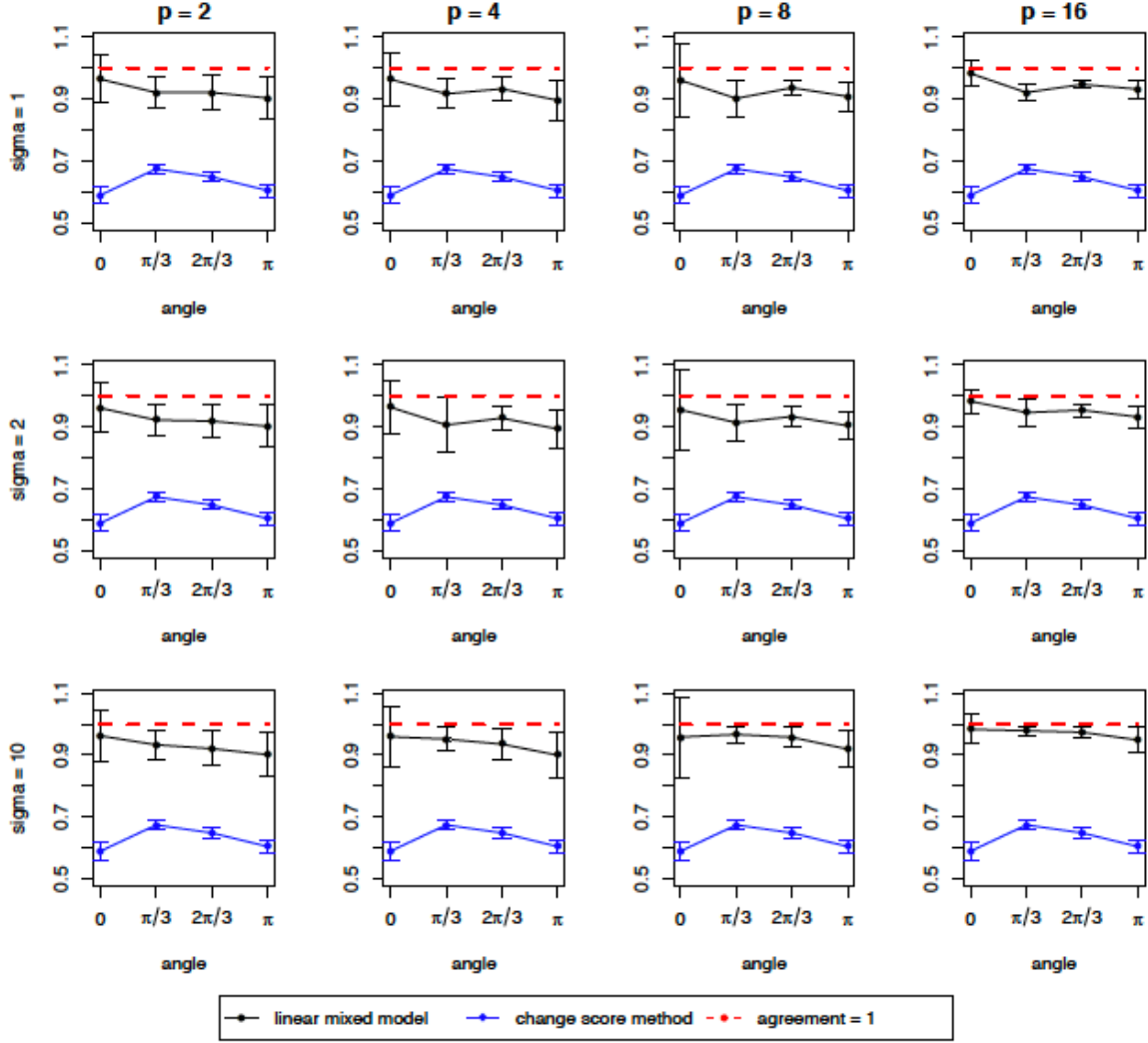
Table 4: Change score

| Dimension | Sigma | Angle | Concordance | | Dimension | Sigma | Angle | Concordance | |
			Mean	SD				Mean	SD
2	1	0	0.590	0.028	8	1	0	0.592	0.030
		60	0.674	0.016			60	0.716	0.017
		120	0.649	0.016			120	0.688	0.018
		180	0.605	0.021			180	0.629	0.017
	2	0	0.589	0.028		2	0	0.592	0.030
		60	0.675	0.016			60	0.715	0.017
		120	0.648	0.016			120	0.688	0.018
		180	0.605	0.020			180	0.629	0.017
	10	0	0.583	0.028		10	0	0.586	0.030
		60	0.666	0.017			60	0.708	0.017
		120	0.641	0.018			120	0.680	0.018

Dimension	Sigma	Angle	Mean	SD	Dimension	Sigma	Angle	Mean	SD
		180	0.599	0.018			180	0.621	0.017
4	1	0	0.590	0.027	16	1	0	0.591	0.022
		60	0.692	0.014			60	0.739	0.014
		120	0.665	0.015			120	0.714	0.015
		180	0.613	0.019			180	0.650	0.016
	2	0	0.590	0.027		2	0	0.591	0.022
		60	0.692	0.014			60	0.739	0.014
		120	0.665	0.015			120	0.713	0.015
		180	0.613	0.019			180	0.649	0.016
	10	0	0.580	0.028		10	0	0.585	0.024
		60	0.684	0.015			60	0.733	0.015
		120	0.656	0.015			120	0.708	0.016
		180	0.605	0.018			180	0.641	0.017

Table 5: Longitudinal decision rule

Dimension	Sigma	Angle	Purity			Cosine similarity		L2norm		Concordance	
			True	Mean	SD	Mean	SD	Mean	SD	Mean	SD
2	1	0	0.86	2.61	0.98	0.963	0.058	0.215	0.168	0.964	0.076
		60	0.09	4.11	1.24	0.949	0.086	0.249	0.202	0.921	0.049
		120	2.78	5.66	1.68	0.985	0.023	0.134	0.112	0.921	0.054
		180	5.46	5.06	1.61	0.981	0.032	0.155	0.118	0.903	0.070
	2	0	0.86	2.97	2.18	0.941	0.084	0.270	0.214	0.961	0.079
		60	0.09	4.46	2.55	0.948	0.086	0.236	0.220	0.923	0.050
		120	2.78	6.18	3.40	0.978	0.030	0.168	0.125	0.919	0.053
		180	5.46	5.56	3.42	0.972	0.041	0.186	0.145	0.903	0.069
	10	0	0.86	490.25	3997.54	0.973	0.086	0.126	0.199	0.962	0.082
		60	0.09	1511.93	8407.62	0.983	0.065	0.097	0.156	0.933	0.049
		120	2.78	1977.48	10567.23	0.991	0.024	0.089	0.101	0.921	0.056
		180	5.46	493.90	3738.82	0.958	0.122	0.159	0.245	0.901	0.073
4	1	0	2.86	3.00	0.85	0.916	0.071	0.374	0.166	0.964	0.086
		60	1.27	5.33	1.16	0.870	0.128	0.458	0.226	0.917	0.047
		120	3.21	8.05	1.87	0.957	0.033	0.272	0.112	0.933	0.037
		180	7.67	7.19	1.87	0.953	0.033	0.286	0.111	0.896	0.064
	2	0	2.86	3.52	1.56	0.886	0.101	0.424	0.222	0.964	0.087
		60	1.27	5.94	2.10	0.851	0.196	0.432	0.336	0.907	0.088
		120	3.21	9.21	3.49	0.939	0.055	0.309	0.165	0.929	0.039
		180	7.67	8.40	3.76	0.930	0.058	0.334	0.169	0.894	0.061
	10	0	2.86	1633.82	9357.03	0.978	0.108	0.088	0.190	0.959	0.096
		60	1.27	1740.05	9491.58	0.991	0.035	0.076	0.112	0.952	0.037
		120	3.21	2118.15	6382.75	0.975	0.088	0.116	0.193	0.935	0.051
		180	7.67	2960.42	12064.43	0.976	0.086	0.105	0.194	0.900	0.073
8	1	0	6.86	3.13	0.99	0.802	0.096	0.609	0.160	0.960	0.118
		60	4.00	7.15	1.26	0.748	0.174	0.671	0.233	0.902	0.060
		120	4.07	12.14	2.59	0.892	0.057	0.450	0.117	0.936	0.026
		180	12.10	10.56	2.75	0.877	0.060	0.480	0.122	0.908	0.047
	2	0	6.86	3.96	2.14	0.784	0.141	0.595	0.283	0.956	0.130
		60	4.00	7.96	2.83	0.797	0.188	0.544	0.333	0.914	0.061

		120	4.07	14.17	5.33	0.879	0.080	0.451	0.198	0.934	0.032
		180	12.10	12.31	5.83	0.852	0.094	0.500	0.215	0.905	0.047
	10	0	6.86	2260.81	11880.16	0.999	0.001	0.033	0.019	0.958	0.131
		60	4.00	4070.68	15848.54	0.994	0.033	0.051	0.101	0.967	0.027
		120	4.07	2489.81	11910.50	0.996	0.027	0.044	0.083	0.957	0.033
		180	12.10	2047.69	10611.78	0.999	0.003	0.030	0.026	0.919	0.058
16	1	0	14.86	3.44	1.10	0.740	0.069	0.714	0.097	0.982	0.043
		60	9.45	11.64	2.57	0.707	0.118	0.748	0.163	0.921	0.027
		120	5.80	21.61	5.32	0.854	0.052	0.531	0.098	0.947	0.014
		180	20.96	18.45	4.81	0.827	0.058	0.579	0.108	0.932	0.031
	2	0	14.86	4.25	2.64	0.804	0.160	0.523	0.347	0.983	0.040
		60	9.45	12.71	4.47	0.841	0.185	0.415	0.384	0.947	0.044
		120	5.80	25.63	11.42	0.890	0.096	0.390	0.260	0.955	0.021
		180	20.96	22.21	11.71	0.846	0.121	0.471	0.294	0.933	0.034
	10	0	14.86	2096.56	9001.68	1.000	0.000	0.020	0.006	0.983	0.047
		60	9.45	781.54	3072.92	1.000	0.000	0.020	0.007	0.979	0.014
		120	5.80	842.96	2920.31	1.000	0.000	0.020	0.007	0.974	0.018
		180	20.96	1179.71	4456.81	1.000	0.000	0.020	0.006	0.949	0.040