For this analysis, we’re going to try to get means and standard errors for each time point by group for some longitudinal data. We’ll use multiple imputation to do this task.

**\*Example - Schizophrenia**

**Data analyzed by Hedeker and Gibbons (1997). A randomized trial for schizophrenia. The trial contained 312 patients received drug therapy and 101 received placebo. The measurements were taken at weeks 0, 1, 3, 6,**

**but some subjects have missing data due to dropout. The outcome of interest is severity of illness (1 = normal, ... , 7 = extremely ill);**

**data** BIOS755.schizo;

input ID Group week severity;

sqrtweek = week\*\*(**0.5**);

r\_severity = round(severity);

datalines;

1103 1 0 5.5

1103 1 1 3.0

1103 1 3 2.5

1103 1 6 4.0

1104 1 0 6.0

1104 1 1 3.0

1104 1 3 1.5

1104 1 6 2.5

………………………………

9316 0 1 6.0

9316 0 3 6.5

9316 0 6 6.0

**run**;

**\* To run the imputation, we first want to change the data from long to wide:**

**\* Going from long to wide (this will only output the outcome and id):**

**proc** **transpose** data=BIOS755.schizo out=schizo\_wide prefix=Y;

by id;

id week;

var severity;

**run**;

**\* Merging the wide outcome data with the other covariate data:**

**data** BIOS755.schizo\_wide;

merge BIOS755.schizo schizo\_wide;

by id;

if first.id;

drop week severity sqrtweek r\_severity \_NAME\_ \_LABEL\_;

**run**;

**proc** **print** data = BIOS755.schizo\_wide (obs=**10**);

**run**;

| **Obs** | **ID** | **Group** | **Y0** | **Y1** | **Y3** | **Y6** |
| --- | --- | --- | --- | --- | --- | --- |
| **1** | 1103 | 1 | 5.5 | 3.0 | 2.5 | 4.0 |
| **2** | 1104 | 1 | 6.0 | 3.0 | 1.5 | 2.5 |
| **3** | 1105 | 1 | 4.0 | 3.0 | 1.0 | . |
| **4** | 1106 | 1 | 3.0 | 1.0 | 1.5 | 1.0 |
| **5** | 1107 | 0 | 5.0 | 5.0 | 5.0 | 5.5 |
| **6** | 1108 | 1 | 6.0 | 6.0 | 3.5 | 4.5 |
| **7** | 1109 | 1 | 4.0 | 2.0 | 2.0 | 2.5 |
| **8** | 1110 | 1 | 4.0 | 4.5 | 4.0 | 3.5 |
| **9** | 1111 | 1 | 5.5 | 5.5 | 5.5 | 2.5 |
| **10** | 1113 | 1 | 4.0 | 2.5 | 4.5 | 3.0 |

**proc** **sort** data = BIOS755.schizo\_wide;

by group;

**run**;

**\* Running the imputation on the wide data;**

**\* First have an imputation model for Y1 (Y0 has no missingness);**

**\* Then we have an imputation model for Y3 given Y0, group, and the imputed Y1;**

**\* Then we have an imputation model for Y6 given Y0, group, and the imputed Y1 & Y3;**

**proc** **mi** data=BIOS755.schizo\_wide NIMPUTE=**10** seed=**8675309** out=MI\_wide;

class group;

fcs nbiter=**20** plots=trace regpmm (Y1 = Y0 group Y0\*group/ k = **8**);

fcs nbiter=**20** plots=trace regpmm (Y3 = Y1 Y0 group Y0\*group Y1\*group/ k = **8**);

fcs nbiter=**20** plots=trace regpmm (Y6 = Y3 Y1 Y0 group Y0\*group Y1\*group Y3\*group/ k = **8**);

var Y0 group Y1 Y3 Y6;

**run**;

Imputation Methods

[DISCRIM](https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.3/statug/statug_mi_syntax05.htm#statug.mi.fcsdiscrimopt)

Specifies the discriminant function method

[LOGISTIC](https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.3/statug/statug_mi_syntax05.htm#statug.mi.fcslogisticopt)

Specifies the logistic regression method

[REG](https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.3/statug/statug_mi_syntax05.htm#statug.mi.fcsregopt)

Specifies the regression method

[REGPMM](https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.3/statug/statug_mi_syntax05.htm#statug.mi.fcsregpmmopt)

Specifies the predictive mean matching method

**ORDER=FREQ | VAR**

specifies the variable ordering in which to impute missing values in the filled-in and imputation phases. The ORDER=FREQ option orders the variables by the descending frequency counts of variables and the ORDER=VAR orders the variables as specified in the VAR statement. The default is ORDER=FREQ.

|  |
| --- |
| The SAS System |

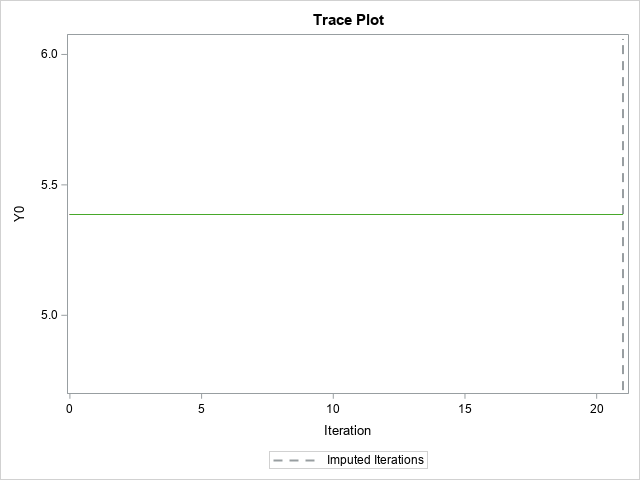
The MI Procedure

| **Model Information** | |
| --- | --- |
| **Data Set** | BIOS755.SCHIZO\_WIDE |
| **Method** | FCS |
| **Number of Imputations** | 10 |
| **Number of Burn-in Iterations** | 20 |
| **Seed for random number generator** | 8675309 |

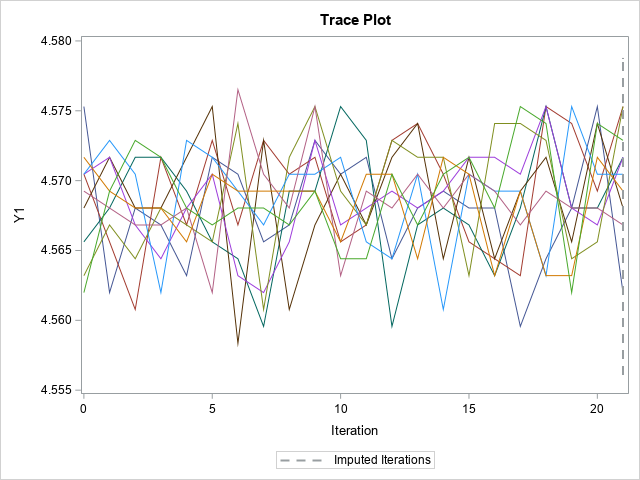
| **FCS Model Specification** | |
| --- | --- |
| **Method** | **Imputed Variables** |
| Regression | Y0 |
| Regression-PMM( K= 8) | Y1 Y3 Y6 |
| Discriminant Function | Group |

| **Missing Data Patterns** | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Group** | **Y0** | **Group** | **Y1** | **Y3** | **Y6** | **Freq** | **Percent** | **Group Means** | | | |
| **Y0** | **Y1** | **Y3** | **Y6** |
| **1** | X | X | X | X | X | 312 | 75.54 | 5.379808 | 4.562821 | 4.092949 | 3.349038 |
| **2** | X | X | X | X | . | 53 | 12.83 | 5.320755 | 4.500000 | 3.641509 | . |
| **3** | X | X | X | . | . | 45 | 10.90 | 5.566667 | 4.744444 | . | . |
| **4** | X | X | . | . | . | 3 | 0.73 | 4.500000 | . | . | . |

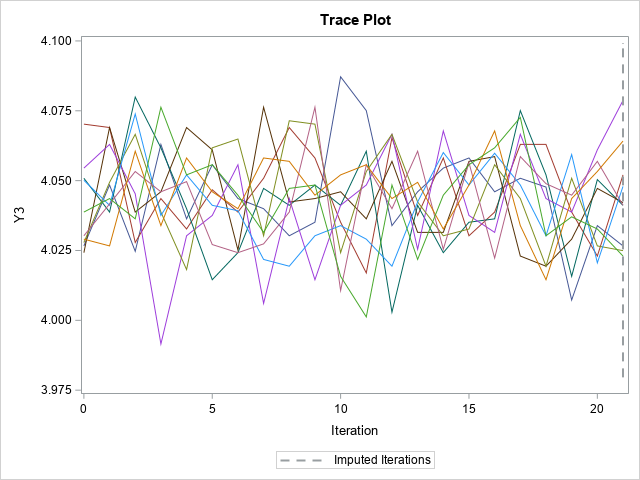
The MI Procedure



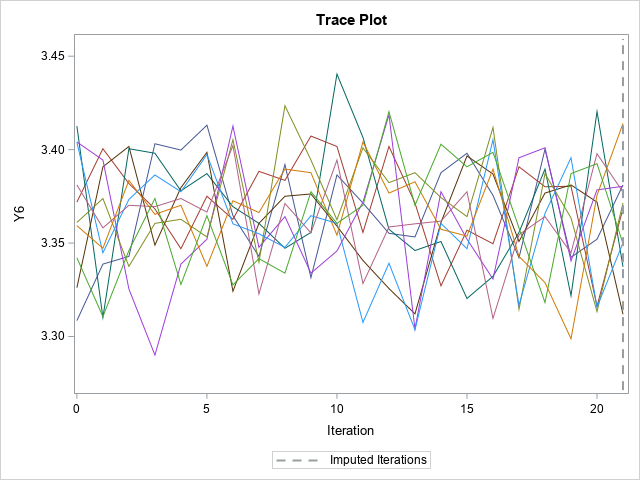
The MI Procedure



The MI Procedure



The MI Procedure



| **Variance Information (10 Imputations)** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Variance** | | | **DF** | **Relative Increase in Variance** | **Fraction Missing Information** | **Relative Efficiency** |
| **Between** | **Within** | **Total** |
| **Y1** | 0.000016432 | 0.003695 | 0.003714 | 407.58 | 0.004891 | 0.004873 | 0.999513 |
| **Y3** | 0.000313 | 0.005299 | 0.005643 | 332.17 | 0.064940 | 0.061755 | 0.993862 |
| **Y6** | 0.000794 | 0.005859 | 0.006733 | 214 | 0.149081 | 0.132977 | 0.986877 |

| **Parameter Estimates (10 Imputations)** | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Mean** | **Std Error** | **95% Confidence Limits** | | **DF** | **Minimum** | **Maximum** | **Mu0** | **t for H0: Mean=Mu0** | **Pr > |t|** |
| **Y1** | 4.570339 | 0.060939 | 4.450545 | 4.690133 | 407.58 | 4.561985 | 4.575303 | 0 | 75.00 | <.0001 |
| **Y3** | 4.044189 | 0.075123 | 3.896412 | 4.191965 | 332.17 | 4.023002 | 4.078692 | 0 | 53.83 | <.0001 |
| **Y6** | 3.364165 | 0.082054 | 3.202427 | 3.525903 | 214 | 3.311864 | 3.414044 | 0 | 41.00 | <.0001 |

**proc** **print** data = MI\_wide (obs=**10**);

**run**;

| **Obs** | **\_Imputation\_** | **ID** | **Group** | **Y0** | **Y1** | **Y3** | **Y6** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **1** | 1 | 1107 | 0 | 5.0 | 5.0 | 5.0 | 5.5 |
| **2** | 1 | 1115 | 0 | 5.5 | 4.5 | 5.5 | 4.5 |
| **3** | 1 | 1124 | 0 | 5.0 | 5.0 | 4.0 | 4.5 |
| **4** | 1 | 1129 | 0 | 6.0 | 4.5 | 3.5 | 2.0 |
| **5** | 1 | 1136 | 0 | 5.0 | 5.5 | 4.5 | 5.5 |
| **6** | 1 | 1140 | 0 | 5.5 | 6.0 | 6.0 | 6.5 |
| **7** | 1 | 1301 | 0 | 5.0 | 2.5 | 2.5 | 3.5 |
| **8** | 1 | 1306 | 0 | 4.5 | 5.0 | 1.5 | 2.5 |
| **9** | 1 | 1309 | 0 | 6.0 | 6.0 | 3.5 | 2.0 |
| **10** | 1 | 1313 | 0 | 6.5 | 6.0 | 6.0 | 6.0 |

**\* Now we’ll do a simple analysis (means and standard errors) for each timepoint based on the imputed data;**

**\* First, we’ll analyze the data by \_imputation\_;**

**proc** **means** data=MI\_wide mean stderr;

by \_imputation\_ group;

var Y0 Y1 Y3 Y6;

output out=MI\_mn mean=mn\_Y0 mn\_Y1 mn\_Y3 mn\_Y6 stderr=SE\_Y0 SE\_Y1 SE\_Y3 SE\_Y6;

**run** ;

**proc** **sort** data=MI\_mn ;

by group \_imputation\_ ;

**run** ;

**\* Second, we’ll use mianalyze to summarize;**

**proc** **mianalyze** data=MI\_mn ;

by group ;

modeleffects mn\_Y0 mn\_Y1 mn\_Y3 mn\_Y6 ;

stderr SE\_Y0 SE\_Y1 SE\_Y3 SE\_Y6;

ods output parameterestimates=outcombine\_1 ;

**run** ;

|  |
| --- |
| The SAS System |

The MIANALYZE Procedure

Group=0

| **Model Information** | |
| --- | --- |
| **Data Set** | WORK.MI\_MN |
| **Number of Imputations** | 10 |

| **Variance Information (10 Imputations)** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Parameter** | **Variance** | | | **DF** | **Relative Increase in Variance** | **Fraction Missing Information** | **Relative Efficiency** |
| **Between** | **Within** | **Total** |
| **mn\_Y0** | 0 | 0.006536 | 0.006536 | . | 0 | . | . |
| **mn\_Y1** | 0 | 0.013068 | 0.013068 | . | 0 | . | . |
| **mn\_Y3** | 0.000465 | 0.013100 | 0.013611 | 6379.7 | 0.039025 | 0.037861 | 0.996228 |
| **mn\_Y6** | 0.008918 | 0.019539 | 0.029349 | 80.563 | 0.502033 | 0.350170 | 0.966168 |

| **Parameter Estimates (10 Imputations)** | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Parameter** | **Estimate** | **Std Error** | **95% Confidence Limits** | | **DF** | **Minimum** | **Maximum** | **Theta0** | **t for H0: Parameter=Theta0** | **Pr > |t|** |
| **mn\_Y0** | 5.373267 | 0.080848 | . | . | . | 5.373267 | 5.373267 | 0 | . | . |
| **mn\_Y1** | 4.990099 | 0.114317 | . | . | . | 4.990099 | 4.990099 | 0 | . | . |
| **mn\_Y3** | 4.862475 | 0.116667 | 4.633768 | 5.091183 | 6379.7 | 4.829703 | 4.897030 | 0 | 41.68 | <.0001 |
| **mn\_Y6** | 4.492574 | 0.171315 | 4.151682 | 4.833466 | 80.563 | 4.358416 | 4.625743 | 0 | 26.22 | <.0001 |

The MIANALYZE Procedure

Group=1

| **Model Information** | |
| --- | --- |
| **Data Set** | WORK.MI\_MN |
| **Number of Imputations** | 10 |

| **Variance Information (10 Imputations)** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Parameter** | **Variance** | | | **DF** | **Relative Increase in Variance** | **Fraction Missing Information** | **Relative Efficiency** |
| **Between** | **Within** | **Total** |
| **mn\_Y0** | 0 | 0.002423 | 0.002423 | . | 0 | . | . |
| **mn\_Y1** | 0.000028793 | 0.004877 | 0.004909 | 216214 | 0.006494 | 0.006461 | 0.999354 |
| **mn\_Y3** | 0.000816 | 0.007005 | 0.007903 | 697.41 | 0.128158 | 0.116130 | 0.988520 |
| **mn\_Y6** | 0.001135 | 0.006478 | 0.007727 | 344.47 | 0.192805 | 0.166465 | 0.983626 |

| **Parameter Estimates (10 Imputations)** | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Parameter** | **Estimate** | **Std Error** | **95% Confidence Limits** | | **DF** | **Minimum** | **Maximum** | **Theta0** | **t for H0: Parameter=Theta0** | **Pr > |t|** |
| **mn\_Y0** | 5.390385 | 0.049222 | . | . | . | 5.390385 | 5.390385 | 0 | . | . |
| **mn\_Y1** | 4.434455 | 0.070064 | 4.297131 | 4.571779 | 216214 | 4.423397 | 4.441026 | 0 | 63.29 | <.0001 |
| **mn\_Y3** | 3.779295 | 0.088898 | 3.604755 | 3.953835 | 697.41 | 3.742628 | 3.832372 | 0 | 42.51 | <.0001 |
| **mn\_Y6** | 2.998878 | 0.087904 | 2.825981 | 3.171775 | 344.47 | 2.952564 | 3.054167 | 0 | 34.12 | <.0001 |

**proc** **print** data=outcombine\_1;

**run**;

| **Obs** | **Group** | **NImpute** | **Parm** | **Estimate** | **StdErr** | **LCLMean** | **UCLMean** | **DF** | **Min** | **Max** | **Theta0** | **tValue** | **Probt** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **1** | 0 | 10 | mn\_Y0 | 5.373267 | 0.080848 | . | . | . | 5.373267 | 5.373267 | 0 | . | . |
| **2** | 0 | 10 | mn\_Y1 | 4.990099 | 0.114317 | . | . | . | 4.990099 | 4.990099 | 0 | . | . |
| **3** | 0 | 10 | mn\_Y3 | 4.862475 | 0.116667 | 4.633768 | 5.091183 | 6379.7 | 4.829703 | 4.897030 | 0 | 41.68 | <.0001 |
| **4** | 0 | 10 | mn\_Y6 | 4.492574 | 0.171315 | 4.151682 | 4.833466 | 80.563 | 4.358416 | 4.625743 | 0 | 26.22 | <.0001 |
| **5** | 1 | 10 | mn\_Y0 | 5.390385 | 0.049222 | . | . | . | 5.390385 | 5.390385 | 0 | . | . |
| **6** | 1 | 10 | mn\_Y1 | 4.434455 | 0.070064 | 4.297131 | 4.571779 | 216214 | 4.423397 | 4.441026 | 0 | 63.29 | <.0001 |
| **7** | 1 | 10 | mn\_Y3 | 3.779295 | 0.088898 | 3.604755 | 3.953835 | 697.41 | 3.742628 | 3.832372 | 0 | 42.51 | <.0001 |
| **8** | 1 | 10 | mn\_Y6 | 2.998878 | 0.087904 | 2.825981 | 3.171775 | 344.47 | 2.952564 | 3.054167 | 0 | 34.12 | <.0001 |

**proc** **means** data = BIOS755.schizo\_wide mean stderr clm alpha=**0.05**;

var Y0 Y1 Y3 Y6;

by group;

**run**;

The MEANS Procedure

Group=0

| **Variable** | **Mean** | **Std Error** | **Lower 95% CL for Mean** | **Upper 95% CL for Mean** |
| --- | --- | --- | --- | --- |
| |  | | --- | | **Y0** | | **Y1** | | **Y3** | | **Y6** | | |  | | --- | | 5.3732673 | | 4.9900990 | | 4.7927711 | | 4.3156250 | | |  | | --- | | 0.0808481 | | 0.1143168 | | 0.1234220 | | 0.1781494 | | |  | | --- | | 5.2128670 | | 4.7632978 | | 4.5472455 | | 3.9596219 | | |  | | --- | | 5.5336677 | | 5.2169002 | | 5.0382967 | | 4.6716281 | |

Group=1

| **Variable** | **Mean** | **Std Error** | **Lower 95% CL for Mean** | **Upper 95% CL for Mean** |
| --- | --- | --- | --- | --- |
| |  | | --- | | **Y0** | | **Y1** | | **Y3** | | **Y6** | | |  | | --- | | 5.3903846 | | 4.4388350 | | 3.8021277 | | 3.0995968 | | |  | | --- | | 0.0492222 | | 0.0703168 | | 0.0867251 | | 0.0892931 | | |  | | --- | | 5.2935341 | | 4.3004728 | | 3.6314143 | | 2.9237238 | | |  | | --- | | 5.4872352 | | 4.5771971 | | 3.9728411 | | 3.2754698 | |