ST 790, Homework 3 Solutions
Spring 2017

1. (a) Here are results using $M = 10$ imputations from `proc mi` and `proc mianalyze` in SAS.

```
data armd; infile "armd.dat";
    input pid line0 lost4 lost12 lost24 lost52 visual0 visual4 visual12 visual24
    visual52 lesion trt;
run;

data armd; set armd;
   if trt = 1 then treat = 0;  *  placebo;
   if trt = 4 then treat = 1;  *  active treatment;
   drop trt line0 lost4 lost12 lost24 lost52 lesion;
run;

proc mi data=armd out=armdout seed=1518971 simple nimpute=10;
   mcmc initial=em chain=single impute=full nbiter=200 niter=100;
   var visual0 visual4 visual12 visual24 visual52;
run;

data armdout_alt; set armdout;
    array vis(5) visual0 visual4 visual12 visual24 visual52;
    do time = 1 to 5;
        visual = vis(time);
        output;
        end;
    drop visual0 visual4 visual12 visual24 visual52;
run;

data armdout_alt; set armdout_alt;
    if time = 1 then week = 0;
    if time = 2 then week = 4;
    if time = 3 then week = 12;
    if time = 4 then week = 24;
    if time = 5 then week = 52;
    drop time;
run;

proc sort data=armdout_alt;
    by _imputation_ pid week;
run;

proc mixed data=armdout_alt asycov covtest method=ml noclprint noitprint;
    class week pid;
    by _imputation_;
    model visual = week week*treat / noint solution covb;
```

```
    repeated week / subject=pid type=un;
    ods output SolutionF=meanparms CovB=meancov CovParms=covparms AsyCov=covcov;
run;

proc mianalyze parms=meanparms covb(effectvar=rowcol)=meancov wcov
    bcov;
    class week;
    modeleffects week week*treat;
run;
```

The MIANALYZE Procedure

Model Information

PARMS Data Set	WORK.MEANPARMS
COVB Data Set	WORK.MEANCOV
Number of Imputations	10

Parameter Estimates

Parameter	week	Estimate	Std Error	95% Confidence Limits	
week	0	55.336134	1.361229	52.66818	58.00409
week	4.000000	54.032935	1.455234	51.18072	56.88515
week	12.000000	52.935760	1.582204	49.83468	56.03683
week	24.000000	49.235479	1.729045	45.84629	52.62467
week	52.000000	43.674754	1.780916	40.18229	47.16722
treat*week	0	-0.757622	1.917097	-4.51506	2.99982
treat*week	4.000000	-2.908678	2.068326	-6.96281	1.14545
treat*week	12.000000	-4.170606	2.256380	-8.59356	0.25235
treat*week	24.000000	-3.648281	2.466492	-8.48423	1.18767
treat*week	52.000000	-4.898175	2.549743	-9.90157	0.10522

Parameter Estimates

Parameter	week	DF	Minimum	Maximum
week	0	217E45	55.336134	55.336134
week	4.000000	398217	53.899386	54.204566
week	12.000000	325206	52.792613	53.093421
week	24.000000	12610	48.612214	49.556411
week	52.000000	2189.2	43.082349	44.354363
treat*week	0	682E45	-0.757622	-0.757622
treat*week	4.000000	17314	-3.360620	-2.360595
treat*week	12.000000	10096	-4.848310	-3.712591
treat*week	24.000000	3417	-4.461658	-2.827102
treat*week	52.000000	1010.2	-5.826761	-3.665325

Parameter Estimates

Parameter	week	Theta0	t for H0: Parameter=Theta0	Pr > |t|
week	0	0	40.65	<.0001
week	4.000000	0	37.13	<.0001
week	12.000000	0	33.46	<.0001
week	24.000000	0	28.48	<.0001
week	52.000000	0	24.52	<.0001
treat*week	0	0	-0.40	0.6927
treat*week	4.000000	0	-1.41	0.1597
treat*week	12.000000	0	-1.85	0.0646
treat*week	24.000000	0	-1.48	0.1392
treat*week	52.000000	0	-1.92	0.0550

Within-Imputation Covariance Matrix

Row	Effect	week	Col1	Col2	Col3
1	week	0	1.852943	1.693159	1.586097
2	week	4.000000	1.693159	2.107638	1.925117

Row	Col4	Col5	Col6	Col7	Col8
1	1.546834	1.232736	-1.852943	-1.693159	-1.586097
2	1.859494	1.506535	-1.693159	-2.107638	-1.925117

Row	Col9	Col10
1	-1.546834	-1.232736
2	-1.859494	-1.506535

Row	Effect	week	Col1	Col2	Col3
3	week	12.000000	1.586097	1.925117	2.490201
4	week	24.000000	1.546834	1.859494	2.229122
5	week	52.000000	1.232736	1.506535	1.910803
6	treat*week	0	-1.852943	-1.693159	-1.586097
7	treat*week	4.000000	-1.693159	-2.107638	-1.925117
8	treat*week	12.000000	-1.586097	-1.925117	-2.490201
9	treat*week	24.000000	-1.546834	-1.859494	-2.229122
10	treat*week	52.000000	-1.232736	-1.506535	-1.910803

Row	Col4	Col5	Col6	Col7	Col8

3	2.229122	1.910803	-1.586097	-1.925117	-2.490201
4	2.909730	2.468178	-1.546834	-1.859494	-2.229122
5	2.468178	2.968304	-1.232736	-1.506535	-1.910803
6	-1.546834	-1.232736	3.675259	3.358332	3.145978
7	-1.859494	-1.506535	3.358332	4.180439	3.818415
8	-2.229122	-1.910803	3.145978	3.818415	4.939241
9	-2.909730	-2.468178	3.068100	3.688252	4.421400
10	-2.468178	-2.968304	2.445097	2.988168	3.790022

Row	Col9	Col10
3	-2.229122	-1.910803
4	-2.909730	-2.468178
5	-2.468178	-2.968304
6	3.068100	2.445097
7	3.688252	2.988168
8	4.421400	3.790022
9	5.771366	4.895560
10	4.895560	5.887544

Between-Imputation Covariance Matrix

Row	Effect	week	Col1	Col2	Col3
1	week	0	1.084208E-23	-8.9132E-14	5.516193E-14
2	week	4.000000	-8.9132E-14	0.009152	0.005159
3	week	12.000000	5.516193E-14	0.005159	0.011972
4	week	24.000000	-3.0131E-14	0.006103	0.017813
5	week	52.000000	-2.49378E-13	-0.005240	0.013442
6	treat*week	0	-1.11153E-23	1.245724E-13	-7.23773E-14
7	treat*week	4.000000	-6.0869E-14	-0.015319	0.005673
8	treat*week	12.000000	1.693196E-13	-0.011294	0.004219
9	treat*week	24.000000	6.189596E-13	-0.009105	0.007651
10	treat*week	52.000000	1.57137E-12	-0.014738	0.005545

Row	Col4	Col5	Col6	Col7	Col8
1	-3.0131E-14	-2.49378E-13	-1.11153E-23	-6.0869E-14	1.693196E-13
2	0.006103	-0.005240	1.245724E-13	-0.015319	-0.011294
3	0.017813	0.013442	-7.23773E-14	0.005673	0.004219
4	0.072607	0.077731	4.921414E-14	0.008221	0.000265
5	0.077731	0.184873	2.259791E-13	0.071228	0.103291
6	4.921414E-14	2.259791E-13	1.213451E-23	-9.49389E-14	-2.30015E-13
7	0.008221	0.071228	-9.49389E-14	0.088667	0.071521
8	0.000265	0.103291	-2.30015E-13	0.071521	0.138191
9	-0.058655	-0.005885	-8.4083E-13	0.066230	0.129024
10	-0.076576	-0.135297	-1.89231E-12	-0.008444	0.032413

Row	Col9	Col10
1	6.189596E-13	1.57137E-12
2	-0.009105	-0.014738
3	0.007651	0.005545
4	-0.058655	-0.076576
5	-0.005885	-0.135297
6	-8.4083E-13	-1.89231E-12
7	0.066230	-0.008444
8	0.129024	0.032413
9	0.283833	0.272394
10	0.272394	0.557860

(b) Here are results using $M = 10$ imputations from R.

```
mi.mv.results <- cbind(mi.mv.mu$est,mi.mv.mu$std.err,mi.mv.mu$lower,mi.mv.mu$upper,
      mi.mv.mu$df,apply(mu.imp,1,min),apply(mu.imp,1,max))
colnames(mi.mv.results) <- c("Est","StdErr","Lower","Upper","DF","Min","Max")
```

```
> mi.mv.results
                     Est   StdErr      Lower      Upper           DF        Min
 week0        55.3361345 1.366923  52.657014 58.0152547 5.692902e+49 55.3361345
 week4        53.9983351 1.470656  51.115891 56.8807794 3.211002e+05 53.8905046
 week12       53.0367659 1.600221  49.900357 56.1731744 1.125988e+05 52.8112682
 week24       49.3658221 1.722016  45.990628 52.7410158 3.914704e+04 49.0577903
 week52       44.0795988 1.740262  40.668617 47.4905809 3.131503e+04 43.8306617
 week0:treat  -0.7576221 1.925117  -4.530781  3.0155373 1.278515e+53 -0.7576221
 week4:treat  -2.8427233 2.078836  -6.917253  1.2318063 5.688410e+04 -3.2633889
 week12:treat -4.2293412 2.250474  -8.640212  0.1815299 2.408761e+05 -4.4519725
 week24:treat -3.9801215 2.431066  -8.745177  0.7849340 2.273137e+04 -4.5386600
 week52:treat -5.3368113 2.501575 -10.241903 -0.4317194 2.832498e+03 -6.2416103
                     Max
 week0        55.3361345
 week4        54.1360658
 week12       53.2369302
 week24       49.6472523
 week52       44.4592236
 week0:treat  -0.7576221
 week4:treat  -2.5403900
 week12:treat -3.9595350
 week24:treat -3.4049473
 week52:treat -4.1863110
```

```
within.cov <- mi.mv.mu$within
between.cov <- mi.mv.mu$between
Rubin.cov <- mi.mv.mu$cov.mat
```

```
> within.cov
                 week0     week4    week12    week24    week52 week0:treat
week0         1.868479  1.720855  1.616087  1.549437  1.229337   -1.868479
week4         1.720855  2.151380  1.963018  1.865953  1.485756   -1.720855
week12        1.616087  1.963018  2.537812  2.242389  1.914263   -1.616087
week24        1.549437  1.865953  2.242389  2.920377  2.482877   -1.549437
week52        1.229337  1.485756  1.914263  2.482877  2.977169   -1.229337
week0:treat  -1.868479 -1.720855 -1.616087 -1.549437 -1.229337    3.706074
week4:treat  -1.720855 -2.151380 -1.963018 -1.865953 -1.485756    3.413266
week12:treat -1.616087 -1.963018 -2.537812 -2.242389 -1.914263    3.205462
week24:treat -1.549437 -1.865953 -2.242389 -2.920377 -2.482877    3.073263
week52:treat -1.229337 -1.485756 -1.914263 -2.482877 -2.977169    2.438354
             week4:treat week12:treat week24:treat week52:treat
week0          -1.720855    -1.616087    -1.549437    -1.229337
week4          -2.151380    -1.963018    -1.865953    -1.485756
week12         -1.963018    -2.537812    -2.242389    -1.914263
week24         -1.865953    -2.242389    -2.920377    -2.482877
week52         -1.485756    -1.914263    -2.482877    -2.977169
week0:treat     3.413266     3.205462     3.073263     2.438354
week4:treat     4.267199     3.893589     3.701063     2.946955
week12:treat    3.893589     5.033677     4.447715     3.796885
week24:treat    3.701063     4.447715     5.792483     4.924715
week52:treat    2.946955     3.796885     4.924715     5.905128

> between.cov
                      week0         week4        week12        week24
week0         6.753827e-25  4.653107e-14  5.783226e-14 -3.086687e-15
week4         4.653107e-14  1.040952e-02  1.096964e-02  4.253583e-03
week12        5.783226e-14  1.096964e-02  2.081236e-02  2.590612e-03
week24       -3.086687e-15  4.253583e-03  2.590612e-03  4.087460e-02
week52       -8.791220e-14 -3.644173e-03 -1.488264e-03  1.169812e-02
week0:treat   6.881751e-26  8.089512e-16 -2.830497e-15 -4.744105e-15
week4:treat  -3.732046e-14 -1.709397e-02 -2.237115e-02 -8.964113e-03
week12:treat -4.880165e-14 -3.915117e-03 -1.591424e-03  9.340653e-03
week24:treat -1.055093e-13 -7.780987e-03 -4.752476e-03 -4.177296e-02
week52:treat  9.146363e-14  2.054139e-02 -1.565576e-03  2.201431e-02
                    week52   week0:treat   week4:treat  week12:treat
week0        -8.791220e-14  6.881751e-26 -3.732046e-14 -4.880165e-14
week4        -3.644173e-03  8.089512e-16 -1.709397e-02 -3.915117e-03
week12       -1.488264e-03 -2.830497e-15 -2.237115e-02 -1.591424e-03
week24        1.169812e-02 -4.744105e-15 -8.964113e-03  9.340653e-03
week52        4.667468e-02 -2.646549e-14 -1.951019e-02  2.454565e-03
week0:treat  -2.646549e-14  2.826766e-26  1.504307e-14 -6.822907e-15
week4:treat  -1.951019e-02  1.504307e-14  4.941666e-02 -1.183270e-03
week12:treat  2.454565e-03 -6.822907e-15 -1.183270e-03  2.814362e-02
week24:treat -2.439624e-02  5.621647e-15  2.465615e-02  5.839424e-03
week52:treat -7.622072e-02  3.616309e-14 -1.496488e-03  3.914533e-02
             week24:treat  week52:treat
```

```
week0        -1.055093e-13  9.146363e-14
week4        -7.780987e-03  2.054139e-02
week12       -4.752476e-03 -1.565576e-03
week24       -4.177296e-02  2.201431e-02
week52       -2.439624e-02 -7.622072e-02
week0:treat   5.621647e-15  3.616309e-14
week4:treat   2.465615e-02 -1.496488e-03
week12:treat  5.839424e-03  3.914533e-02
week24:treat  1.069078e-01  3.270692e-02
week52:treat  3.270692e-02  3.206791e-01

> Rubin.cov
                   week0     week4    week12    week24    week52 week0:treat
week0          1.868479  1.720855  1.616087  1.549437  1.229337   -1.868479
week4          1.720855  2.162830  1.975085  1.870632  1.481748   -1.720855
week12         1.616087  1.975085  2.560706  2.245239  1.912626   -1.616087
week24         1.549437  1.870632  2.245239  2.965339  2.495745   -1.549437
week52         1.229337  1.481748  1.912626  2.495745  3.028511   -1.229337
week0:treat   -1.868479 -1.720855 -1.616087 -1.549437 -1.229337    3.706074
week4:treat   -1.720855 -2.170183 -1.987626 -1.875813 -1.507218    3.413266
week12:treat  -1.616087 -1.967325 -2.539563 -2.232115 -1.911563    3.205462
week24:treat  -1.549437 -1.874512 -2.247617 -2.966327 -2.509713    3.073263
week52:treat  -1.229337 -1.463161 -1.915985 -2.458661 -3.061011    2.438354
              week4:treat week12:treat week24:treat week52:treat
week0           -1.720855    -1.616087    -1.549437    -1.229337
week4           -2.170183    -1.967325    -1.874512    -1.463161
week12          -1.987626    -2.539563    -2.247617    -1.915985
week24          -1.875813    -2.232115    -2.966327    -2.458661
week52          -1.507218    -1.911563    -2.509713    -3.061011
week0:treat      3.413266     3.205462     3.073263     2.438354
week4:treat      4.321557     3.892288     3.728185     2.945309
week12:treat     3.892288     5.064635     4.454138     3.839945
week24:treat     3.728185     4.454138     5.910082     4.960692
week52:treat     2.945309     3.839945     4.960692     6.257875
```

(c) The results are similar but somewhat different (although still within SE of each other). With $M = 10$ imputations, this does not seem unexpected.

(d) For my results, the inferences on β_5 are qualitatively similar. Here, interestingly, we get a test statistic of -1.92 with p-value 0.055 in SAS and -2.13 with p-value 0.032.

(e) This is omitted - you likely found that the results within a software imputation didn't change very much from $M = 10$ to 100, and any differences between SAS and R were similar for $M = 10$ and 100. With $M = 100$, one would hope that presumably identical implementations would give results that are fairly close. For me, within an implementation (SAS or R), results did not change in any appreciable way from $M = 10$ to 100 but the small differences across SAS and R persisted. This might suggest different conventions and so on in the "guts" of the implementations.

2. (a) (a) Here are results using $M = 10$ imputations from `proc mi` and `proc mianalyze` in SAS.

```
data analg; infile "analgesic.dat";
    input pid age weight genhlth physfct gsa0 gsa1 gsa2 gsa3 gsa4
                gsabin0 gsabin1 gsabin2 gsabin3 gsabin4;
run;

proc mi data=analg nimpute=10 out=analgbinout seed=2382352;
    class gsabin0 gsabin1 gsabin2 gsabin3 gsabin4;
     var age weight genhlth physfct gsabin0 gsabin1 gsabin2 gsabin3 gsabin4;
    fcs reg(age);
    fcs reg(weight);
    fcs reg(genhlth);
    fcs reg(physfct);
    fcs logistic(gsabin0);
    fcs logistic(gsabin1);
    fcs logistic(gsabin2);
    fcs logistic(gsabin3);
    fcs logistic(gsabin4);
run;

data analg0_alt;  set analgbinout;
    array y(5) gsabin0 gsabin1 gsabin2 gsabin3 gsabin4;
    do j = 1 to 5;
        gsa = y(j);
        time=j-1;
        time2=time*time;
        timecls=time;
        output;
        end;
    drop j gsabin0 gsabin1 gsabin2 gsabin3 gsabin4 gsa0 gsa1 gsa2 gsa3 gsa4;
run;

proc sort data=analg0_alt;
    by _imputation_ pid time;
run;

proc genmod data=analg0_alt descending;
    by _imputation_;
    class pid timecls;
    model gsa = time genhlth physfct / dist=bin link=logit;
    repeated subject=pid / type=un withinsubject=timecls modelse mcovb;
    ods output GEEModPEst=gmparms ParmInfo=gmpinfo GEENCov=gmcovb;
run;

proc mianalyze parms=gmparms covb=gmcovb parminfo=gmpinfo wcov bcov tcov;
    modeleffects Intercept time genhlth physfct;
run;
```

The MIANALYZE Procedure

Model Information

PARMS Data Set	WORK.GMPARMS
PARMINFO Data Set	WORK.GMPINFO
COVB Data Set	WORK.GMCOVB
Number of Imputations	10

Parameter Estimates

Parameter	Estimate	Std Error	95% Confidence Limits		DF
Intercept	0.756800	0.200613	0.36133	1.152272	210.1
time	0.081888	0.051344	-0.02160	0.185377	43.82

Parameter Estimates

Parameter	Minimum	Maximum
Intercept	0.632616	0.918420
time	0.047677	0.156592

Parameter Estimates

Parameter	Theta0	t for H0: Parameter=Theta0	Pr > |t|
Intercept	0	3.77	0.0002
time	0	1.59	0.1179

Parameter Estimates

Parameter	Estimate	Std Error	95% Confidence Limits		DF
genhlth	0.007984	0.004090	-0.00010	0.016063	155.18
physfct	0.003378	0.003637	-0.00379	0.010546	219.45

Parameter Estimates

Parameter	Minimum	Maximum
genhlth	0.004951	0.010380
physfct	0.000826	0.005633

Parameter Estimates

Parameter	Theta0	t for H0: Parameter=Theta0	Pr > |t|
genhlth	0	1.95	0.0528
physfct	0	0.93	0.3541

Within-Imputation Covariance Matrix

	Intercept	time	genhlth	physfct
Intercept	0.0319159688	-.0022554851	-.0004260930	-.0002251995
time	-.0022554851	0.0014414646	0.0000008499	0.0000003358
genhlth	-.0004260930	0.0000008499	0.0000127011	-.0000022917
physfct	-.0002251995	0.0000003358	-.0000022917	0.0000105515

Between-Imputation Covariance Matrix

	Intercept	time	genhlth	physfct
Intercept	0.0075724260	-.0000663333	-.0001531920	-.0000128363
time	-.0000663333	0.0010860849	0.0000070081	-.0000236586
genhlth	-.0001531920	0.0000070081	0.0000036628	-.0000008114
physfct	-.0000128363	-.0000236586	-.0000008114	0.0000024358

(b) And here is what I got for $M = 10$ using `mice` in R. These results are somewhat different.

```
> mi.mv.results
                    Est      StdErr         Lower      Upper          DF         Min
(Intercept) 0.773572174 0.191080525  0.3981392020 1.14900515  492.81852 0.649179633 0.8537
time        0.095679218 0.058263748 -0.0235251069 0.21488354   28.76897 0.027023179 0.1366
genhlth     0.006550274 0.003630838 -0.0005703143 0.01367086 2015.28768 0.004886382 0.0079
physfct     0.004477205 0.003575131 -0.0025583144 0.01151272  299.92020 0.002139133 0.0070
> within.cov
              (Intercept)          time       genhlth       physfct
(Intercept)  0.0315776369 -2.407777e-03 -4.125485e-04 -2.211440e-04
time        -0.0024077774  1.495966e-03  8.118019e-07  5.782102e-07
genhlth     -0.0004125485  8.118019e-07  1.230200e-05 -2.342615e-06
physfct     -0.0002211440  5.782102e-07 -2.342615e-06  1.056744e-05
> between.cov
              (Intercept)          time       genhlth       physfct
(Intercept)  4.485573e-03 -1.934406e-03 -3.381541e-05 -5.558682e-05
time        -1.934406e-03  1.726089e-03  7.343719e-07  2.983133e-05
genhlth     -3.381541e-05  7.343719e-07  8.008916e-07 -2.555768e-07
physfct     -5.558682e-05  2.983133e-05 -2.555768e-07  2.012842e-06
> Rubin.cov
              (Intercept)          time       genhlth       physfct
(Intercept)  0.0365117670 -4.535624e-03 -4.497454e-04 -2.822895e-04
```

```
time         -0.0045356244  3.394664e-03  1.619611e-06  3.339268e-05
genhlth      -0.0004497454  1.619611e-06  1.318298e-05 -2.623750e-06
physfct      -0.0002822895  3.339268e-05 -2.623750e-06  1.278156e-05
```

(c) You may have gotten pretty similar results, or you may have found differences. This is probably because the results are pretty unstable for this data set with $M = 10$ imputations.

(d) You probably found qualitative differences in the inferences between SAS and R.

(e) You likely found that results changed when moving from $M = 10$ to $M = 100$. With $M = 100$, there are still differences between SAS and R, suggesting that there are likely differences in the way the methods are implemented and in the conventions used (which are impossible to discern from the documentation).

(f) You probably concluded that the imputation strategy does make a difference. This underscores the concerns we discussed in class regarding sensitivity of inferences to implementation choices.