samon 2.0 - the software

May 24, 2016

Two Software choices

1. R

   samon library
   functions with pass to C code

2. SAS

   procedures and macros

Background

- Randomized study with outcome measurements taken at fixed time-points
- Monotone missing data pattern
- Interest is in a comparison of treatment arm means at the last scheduled time-point
- Outcomes are coded as positive integers
- Missing values are coded as -1
- Rows indicate individuals and columns indicate time-points
- Data at the first time-point (the baseline) is never missing

Background

$$time - point3$$

$$\downarrow$$

$$subjects \left\{ \begin{array}{cccccccc}
82 & 88 & 81 & -1 & -1 & -1 & -1 & -1 \\
71 & 75 & 69 & 66 & 62 & 58 & 51 & 48 \\
62 & 63 & -1 & 55 & 61 & 66 & 68 & -1 \\
72 & 63 & -1 & -1 & 62 & 44 & 55 & -1 \\
83 & 62 & 74 & 67 & -1 & -1 & -1 & -1 \\
88 & 92 & 99 & 70 & -1 & -1 & -1 & -1 \\
66 & 71 & 71 & 71 & 75 & 75 & 71 & 71 \\
90 & 88 & 88 & 88 & 77 & -1 & -1 & -1 \\
88 & 91 & 92 & 91 & 95 & 90 & 88 & -1 \\
-1 & 52 & 33 & 99 & 87 & 88 & -1 & -1
\end{array} \right.$$

Case Study: Quetiapine Fumarate

- Among Bipolar I subjects
- Subjects scheduled for 3 assessements at days 1, 29 and 57.
- Primary efficiacy variable: QLESSF score (Quality of Life Enjoyment and Satisfaction – Short Form)
- Two treatment groups considered here: placebo (treatment 1) with quetiapine fumarate 600mg (treatment 2).

QLESSF Analysis

- Descriptive analysis for each treatment group.
- Check the monotone missing condition.
- Decide on a sensitivity function.
- Use the SAMON procedure to estimate means.
- Examining the results.

The QLESSF Dataset

```
     Obs      V1      V2      V3


       1      36      56      57
       2      32      32      31
       3      25      42      27
       4      48      52      54
       5      27      40      40
       6      38      43      -1
       7      35      44      39
       8      24      26      28
       9      25      30      -1
      10      25      28      -1
      11      40      45      45
      12      14      16      -1
      13      34      37      -1
      14      27      25      -1
```

The samonDataCheck macro

- The samonDataCheck macro can be used to check data to ensure it is in samon canonical form.

%samonDataCheck	
(	
data =	input dataset
vars =	variable list (in time order)
out =	output data
stats =	output statistics dataset
mpattern =	missing pattern counts dataset
);	

samonDataCheck

```
> # R version of samonDataCheck is a
> # function of the same name
>
> # Check data
> chkI1  <- samonDataCheck( trtI1  )
>
> chkI2  <- samonDataCheck( trtI2  )
```

samonDataCheck

```
Samon Data Check:
--------------------------------
Number of time-points:         3
Number of subjects:          116
Minimum value:                14
Maximum value:                63

No Samon problems found

Missing Patterns:
               N   proportion
*__   :       10       0.0862
**_   :       41       0.3534
***   :       65       0.5603
```

samonDataCheck

```
Samon Data Check:
--------------------------------
Number of time-points:         3
Number of subjects:          118
Minimum value:                15
Maximum value:                67

No Samon problems found

Missing Patterns:
               N   proportion
*__  :        13       0.1102
**_  :        37       0.3136
***  :        68       0.5763
```

Two natural questions

1. If an individual is on-study at time $t-1$, what is the probability of them staying on-study at time t? This probability may depend on the value Y_{t-1}. This leads to our first model:

   $$Prob[R_t = 1 \mid R_{t-1} = 1, Y_{t-1} = y] \sim smooth(y; \sigma)$$

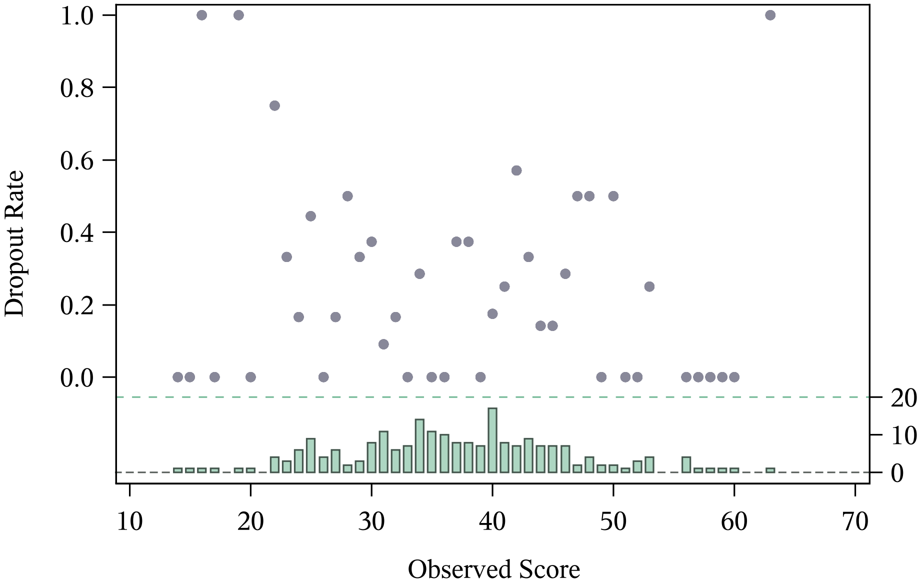
   The smooth function of y depends on a single smoothing parameter σ.

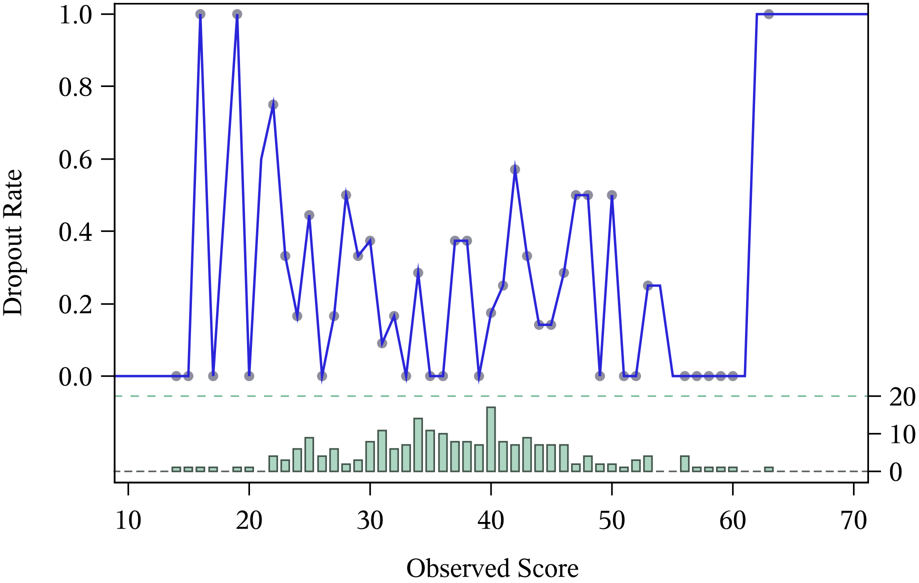
Two natural questions

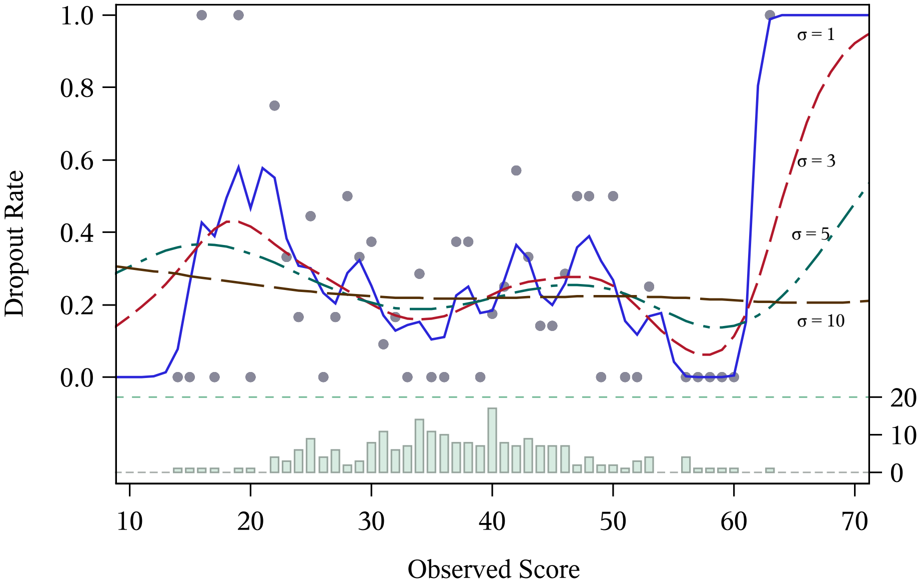
2. If individuals are on-study at time $t-1$ and remain on-study at time t, what is the distribution of their y values at time t? Again this distribution may depend on the value Y_{t-1}. This leads to our second model:

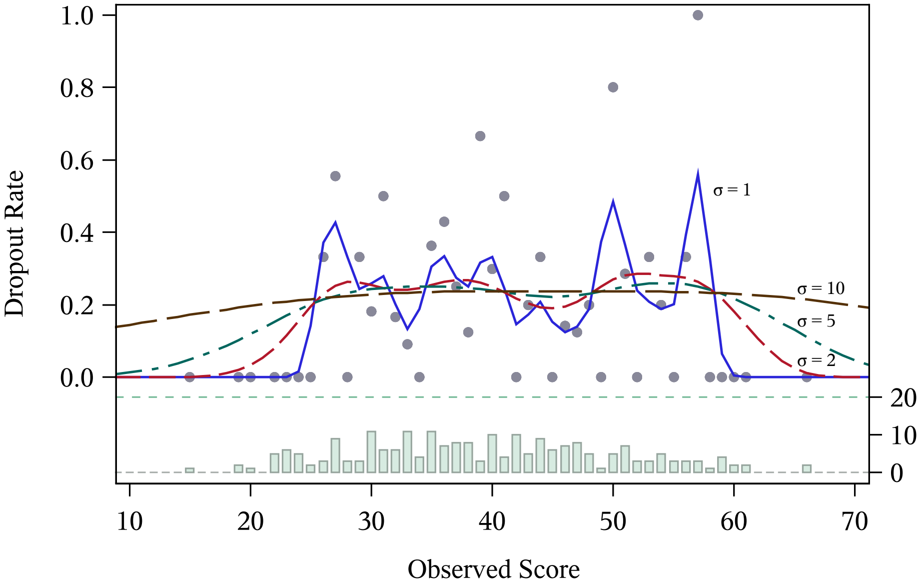
   $h[y' \mid R_t = 1, Y_{t-1} = y] \sim \mathit{smooth}(y;\sigma)$

   The smooth function of y depends on a single smoothing parameter σ.


1.0
0.8
0.6
0.4
0.2
0.0
Dropout Rate
20
10
0
10
20
30
40
50
60
70
Observed Score


1.0
0.8
0.6
0.4
0.2
0.0
Dropout Rate
20
10
0
10
20
30
40
50
60
70
Observed Score


1.0
0.8
0.6
0.4
0.2
0.0
Dropout Rate
σ = 1
σ = 3
σ = 5
σ = 10
20
10
0
10
20
30
40
50
60
70
Observed Score


1.0
0.8
0.6
0.4
0.2
0.0
Dropout Rate
σ = 1
σ = 10
σ = 5
σ = 2
20
10
0
10
20
30
40
50
60
70
Observed Score

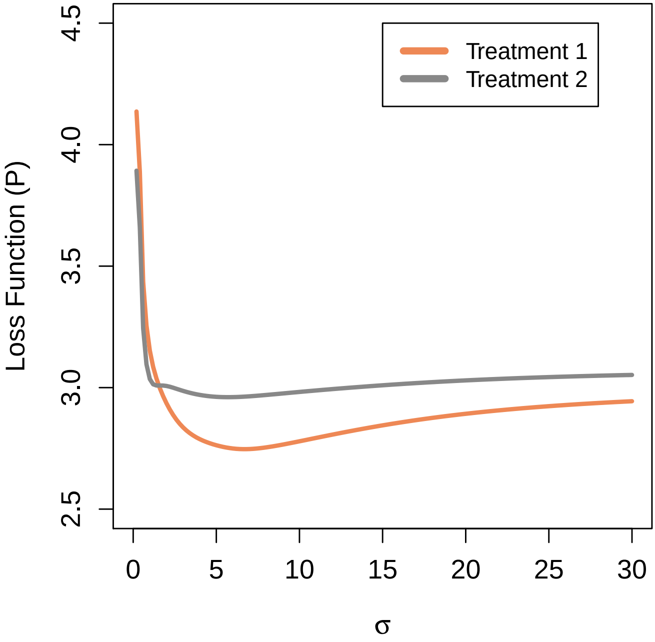
The SAMONEV Procedure

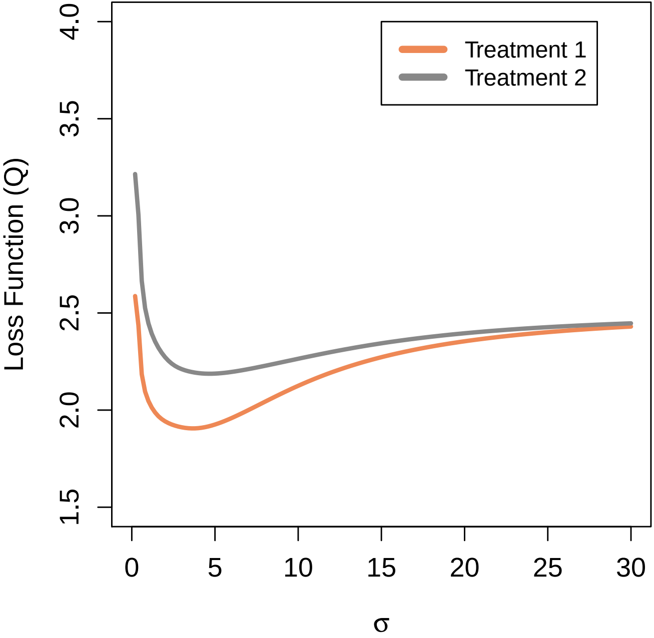
The SAMONEV procedure computes the loss function for a range of σ.

samonev	
data =	Input dataset
out =	Output dataset
npart =	Number of partitions
lowSigma =	low value of sigma
highSigma =	high value of sigma
incSigma =	increment sigma
var varlist	list of variables in time order

The SAMONEV Procedure

```
proc samonev
   data      = quet1
   out       = ev1
   Npart     = 10
   incSigma  = 0.1
   lowSigma  = 0.2
   highSigma = 30;

   var v1 - v3;
run;
```


Treatment 1
Treatment 2
Loss Function (P)
2.5
3.0
3.5
4.0
4.5
0
5
10
15
20
25
30
σ


Treatment 1
Treatment 2
Loss Function (Q)
1.5
2.0
2.5
3.0
3.5
4.0
0
5
10
15
20
25
30
σ

```
* Finding optimal Sigma_p and Sigma_q.;
* -----------------------------------------;
proc samon data   = quet1
           out    = samon1
           Npart  = 10
           pinit  = 6.0
           phigh  = 50.0
           qinit  = 4.0
           qhigh  = 50;
  var v1 - v3;
run;

%split( data = samon1, tag = 1 );
```

Treatment 1

```
proc print data = PM1 noobs;
run;

  rc    Niter      Sigma        loss
   2        3     6.6918      2.7468
```

```
proc print data = QM1 noobs;
run;

  rc    Niter      Sigma        loss
   2        6     3.6771      1.9057
```

Treatment 2

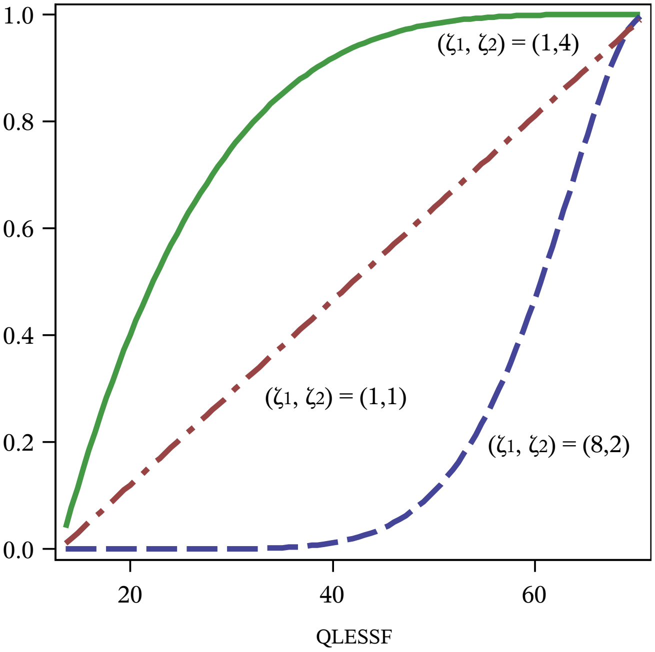
```
proc print data = PM2 noobs;
run;

  rc     Niter      Sigma        loss
   2         3     5.6938      2.9607


proc print data = QM2 noobs;
run;

  rc     Niter      Sigma        loss
   2         3     4.6704      2.1872
```

Sensitivity Analysis

- Within samon the sensitivity bias function is the cumulative function of the beta distribution, a flexible function with bounded support.
- This together with the sensitivity analysis parameter α provides the mechanism by which we measure the sensitivity of the results to informative drop-out.
- $\alpha = 0$ is missing at random
- α quantifies the influence of Y_{t+1} on the decision to drop-out between t and $t + 1$.

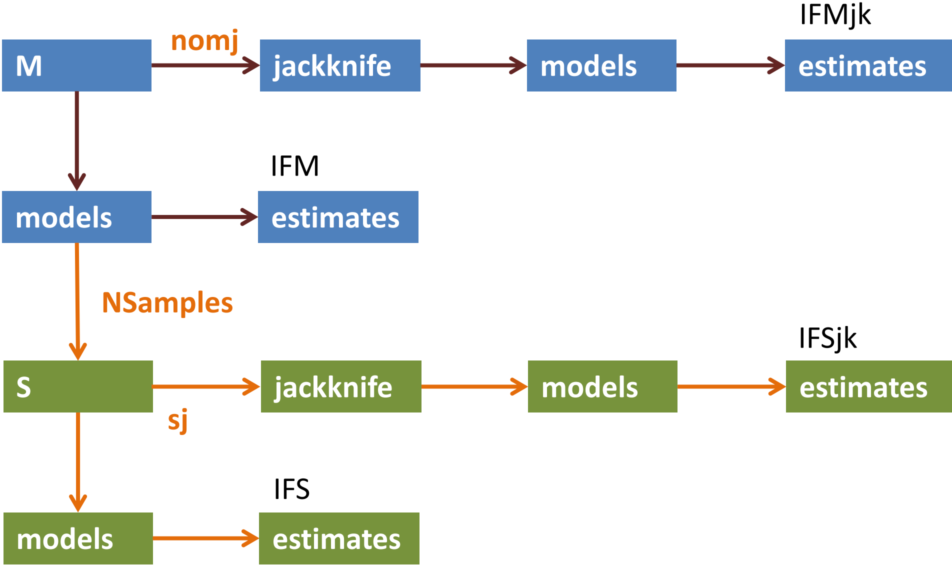
Sensitivity Analysis

- The cumulative beta function is defined on the interval (0,1) and in order to use it as the sensitivity bias function we need to map the range of our data into (0,1).
- In the case of QLESSF scores the data are limited to the range 13 and 71.
- We take the parameters for the cumulative beta function ζ_1 and ζ_2 to be 1.


1.0
0.8
0.6
0.4
0.2
0.0
20
40
60
QLESSF
$(\zeta_1, \zeta_2) = (1,4)$
$(\zeta_1, \zeta_2) = (1,1)$
$(\zeta_1, \zeta_2) = (8,2)$

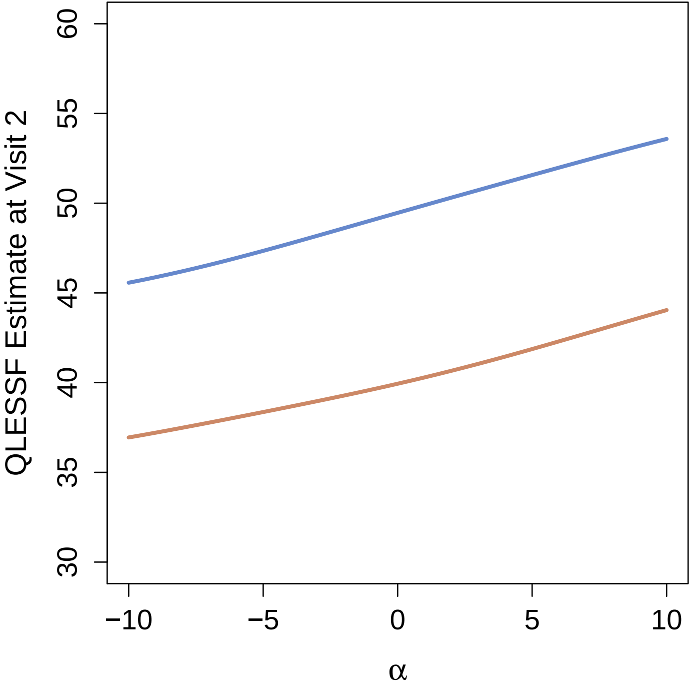
The SAMON procedure

samon	
data =	Input dataset
out =	Output dataset
npart =	Number of partitions
pinit =	initial value for smoothing parameter sigma p
phigh =	Highest value for smoothing parameter sigma p
qinit =	initial value for smoothing parameter sigma q
qhigh =	Highest value for smoothing parameter sigma q
lb =	lower bound of data
ub =	upper bound of data
zeta1 =	parameter for cumulative beta distribution
zeta2 =	parameter for cumulative beta distribution
nsamples =	Number of bootstrap samples
seed0 =	Seed to pass to random number generator
sj	compute jackknifes for each bootstrap sample
nomj	suppress jackknife computation for the main dataset
var varlist	list of variables in time order
sensp senslist	list of sensitivity parameters


IFMjk
M
nomj
jackknife
models
estimates
IFM
models
estimates
NSamples
IFSjk
S
jackknife
models
estimates
sj
IFS
models
estimates

```
proc samon  data = quet1 out = samon1
    Npart = 10
    Pinit =  6.0    PHigh = 50.0
    Qinit =  4.0    QHigh = 50.0

    lb    =   13    ub    =  71
    zeta1 =  1.0    zeta2 = 1.0
    nomj            nsamples = 0 ;

  var v1 - v3;
  sensp -10 to 10 by 1;
run;
%split( data = samon1, tag = 1 );
proc print data = IFM1 noobs;
  var alpha AEst AVar IFEst IFVar;
run;
```

```
alpha      AEst        AVar       IFEst      IFVar
  -10   36.6909     0.08754     36.9454     1.3374
   -9   36.9556     0.09050     37.2108     1.3438
   -8   37.2402     0.09354     37.4881     1.3497
   -7   37.5435     0.09653     37.7740     1.3547
   -6   37.8641     0.09934     38.0654     1.3585
   -5   38.2007     0.10185     38.3611     1.3609
   -4   38.5526     0.10395     38.6609     1.3625
   -3   38.9189     0.10556     38.9660     1.3641
   -2   39.2993     0.10660     39.2787     1.3668
   -1   39.6935     0.10701     39.6020     1.3715
    0   40.1010     0.10678     39.9386     1.3792
    1   40.5210     0.10590     40.2911     1.3898
    2   40.9517     0.10436     40.6609     1.4026
    3   41.3907     0.10222     41.0484     1.4160
    4   41.8343     0.09950     41.4525     1.4278
    5   42.2785     0.09626     41.8710     1.4356
    6   42.7186     0.09252     42.3006     1.4374
    7   43.1501     0.08832     42.7372     1.4320
    8   43.5690     0.08370     43.1760     1.4190
    9   43.9724     0.07871     43.6127     1.3989
```


QLESSF Estimate at Visit 2
30
35
40
45
50
55
60
−10
−5
0
5
10
α

- Use bootstrap with jackknife to compute confidence intervals for IF estimates.
- The NSamples arugment controls the number of bootstraps to make.
- The flags mj and sj control whether jackknifes are performed on the main (input) data and the bootstrap samples respectively.
- For a small dataset with 100 individuals, 1,000 bootstaps each with bootsrap estimates on 50 sensitivity parameters gives rise to 50 x 100 x 1000 = 5 million estimates.

```
proc samon data = quet1 out = samon1
    Npart    = 10

    Pinit    =  6.0    PHigh = 50.0
    QInit    =  4.0    QHigh = 50.0
    lb       = 13      ub    = 71
    zeta1    = 1.0     zeta2 = 1.0
    NSamples = 500     seed0 = 81881
    sj;
  var v1-v3;
  sensp -20 to 20 by 1;
run;
```

macro	description
samonCombine	Combines results from multiple runs of proc samon
samonSummary	Summarizes samon results. Combines bootstrap and jackknife results to produce confidence intervals
samonDifferenceSummary	Computes effect difference with confidence interval from a pair of samonSummary objects.
samonCrossSummary	Computes the effect difference with confidence interval for each pair of sensitivity paramaters α.
samonECompleterStatus	Computes the difference in the expected value of non-completers and completers

samonCombine macro

samonCombine	combines samon results into one dataset
(	
inlib =	input libref
stem = results	file name stem
connect = _	name connector
partfrom = 1	parts start at 1
partto = 100	to 100
partform = z5	format to use on partno
outlib =	output libref
)	

samonSummary macro

samonSummary	computes summary of samon object
(	
data =	input dataset to summarize
out =	summary of main data
sampSummary =	summary of parametric bootstrap samples
)	

samonDifferenceSummary macro

samonDifferenceSummary	Difference in two trials
(	
IFM1 =	main results from samonSummary trt 1
sampIF1 =	sample results from samonSummary trt 1
IFM2 =	main results from samonSummary trt 2
sampIF2 =	sample results from samonSummary trt 2
out =	summary of difference
)	

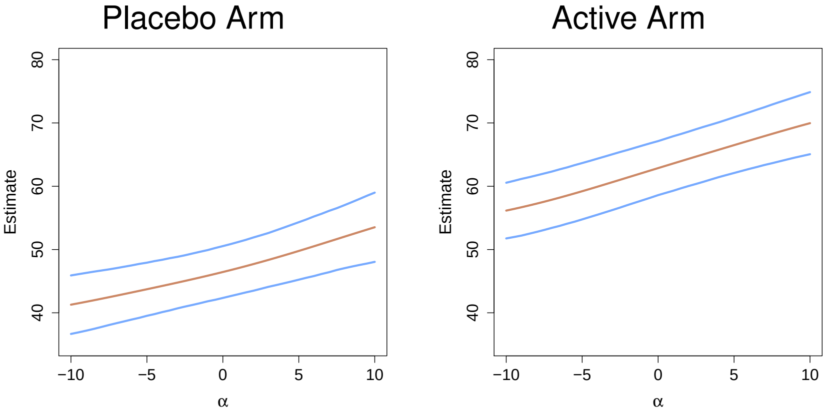
samonCrossSummary macro

samonCrossSummary	Difference in two trials for all pairs of sensitivity parameter
(	
IFM1 =	main results from samonSummary trt 1
sampIF1 =	sample results from samonSummary trt 1
IFM2 =	main results from samonSummary trt 2
sampIF2 =	sample results from samonSummary trt 2
out =	summary of difference
)	

```
%samonSummary(
  data     = results.results1,
  out      = data.Summary1,
  sampout  = data.sampSummary1
  );
proc print data=data.Summary1;
  var alpha IFEst lb ub;
run;
```

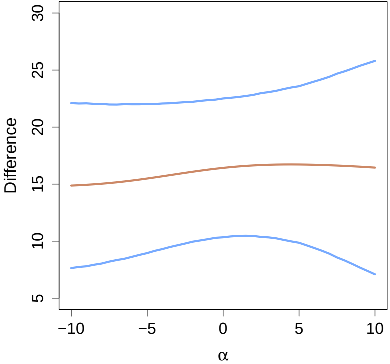
```
alpha     IFEst        lb        ub
  -10   36.9454   34.2645   39.6263
   -9   37.2108   34.5635   39.8580
   -8   37.4881   34.9005   40.0757
   -7   37.7740   35.2496   40.2984
   -6   38.0654   35.5840   40.5468
   -5   38.3611   35.9267   40.7955
   -4   38.6609   36.2627   41.0590
   -3   38.9660   36.6038   41.3281
   -2   39.2787   36.9218   41.6356
   -1   39.6020   37.2546   41.9493
    0   39.9386   37.5628   42.3144
    1   40.2911   37.8957   42.6866
    2   40.6609   38.2211   43.1007
    3   41.0484   38.5820   43.5148
    4   41.4525   38.9005   44.0045
    5   41.8710   39.2378   44.5043
    6   42.3006   39.5695   45.0318
    7   42.7372   39.9254   45.5490
    8   43.1760   40.2880   46.0641
    9   43.6127   40.5862   46.6392
```

Estimated QLESSF score at visit 3



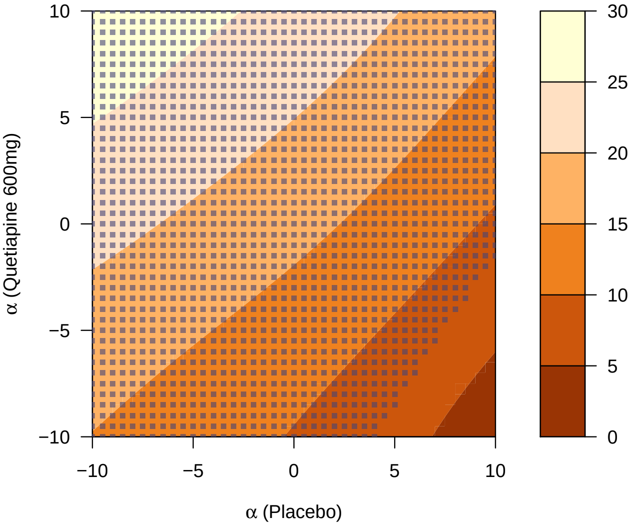
Estimated QLESSF score at visit 3

Difference (active - placebo)



Another useful plot is a surface plot of the difference in the estimated mean value in the two treatment groups given as a function of the two alpha parameters. We use the samonCrossSummary function to compute the difference in estimates for each pair of alpha. The plotting is done with the filled.contour function.

```
%samonCrossSummary(
    IFM1        = data.Summary1,
    sampIF1     = data.sampSummary1,
    IFM2        = data.Summary2,
    sampIF2     = data.sampSummary2,
    out         = data.Cross
    );
```


10
5
0
−5
−10
−10
−5
0
5
10
α (Quetiapine 600mg)
α (Placebo)
30
25
20
15
10
5
0