modelAverage Hands-on

Yasunori Aoki

0: preparation

A standalone software “modelAverage” is available at www.bluetree.me download it and check to see if this software can be started on your computer. Also the Qt source code is available from the website.

1: bootstrap

The software modelAverage essentially combines and summarizes the result of PsN bootstrap analysis given as raw_results files. Before using "modelAverage" you need to run the bootstrap analyses of the multiple candidate models to obtain these raw_results files.

exercise 0
Write a shell script (unix command) to run the bootstrap analyses of the following candidate models:

model0_placebo.mod
model1_linear.mod
model2_loglinear.mod
model3_emax.mod
model4_sigmoidal.mod

You may assume PsN and NONMEM have already been installed and all bootstrap run successfully.

(bonus exercise)
Write the shell-script in a way that it collects all generated raw_results files in a directory so that they can be read easily by modelAverage.

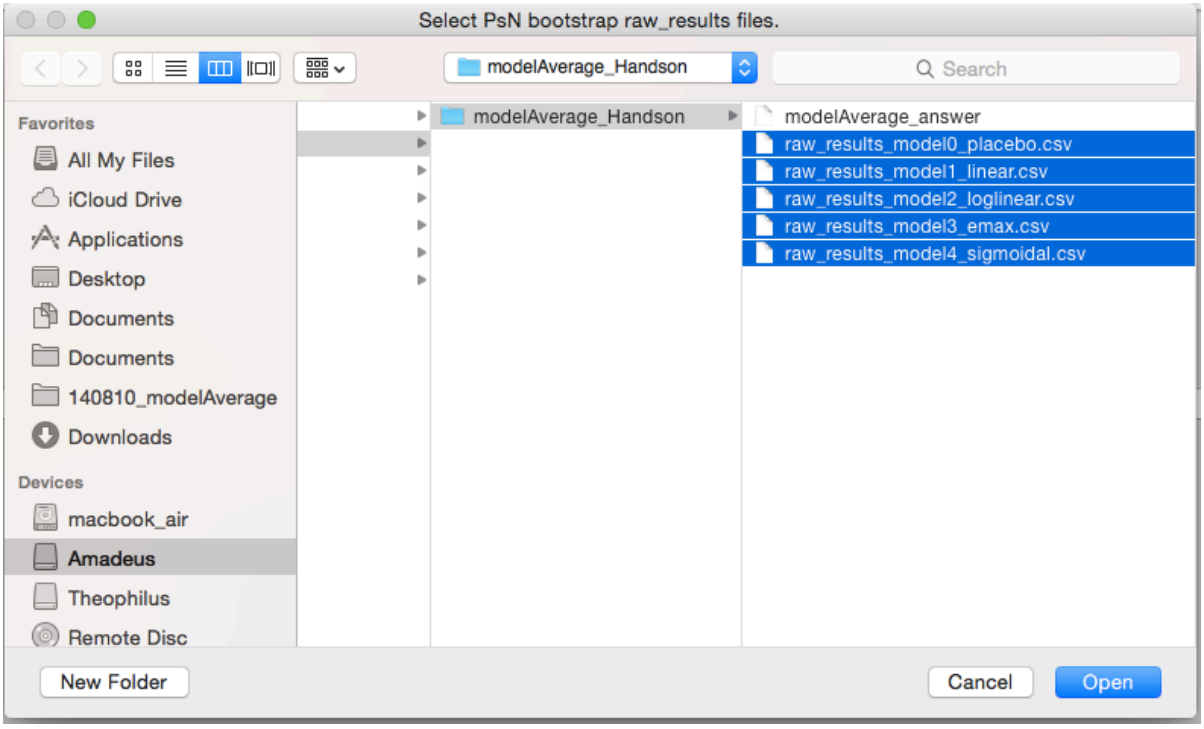
(hint)
It is very important to keep in mind that the set of bootstrap data created by PsN bootstrap must be the same across all the candidate models.

2: Import Raw_Results files

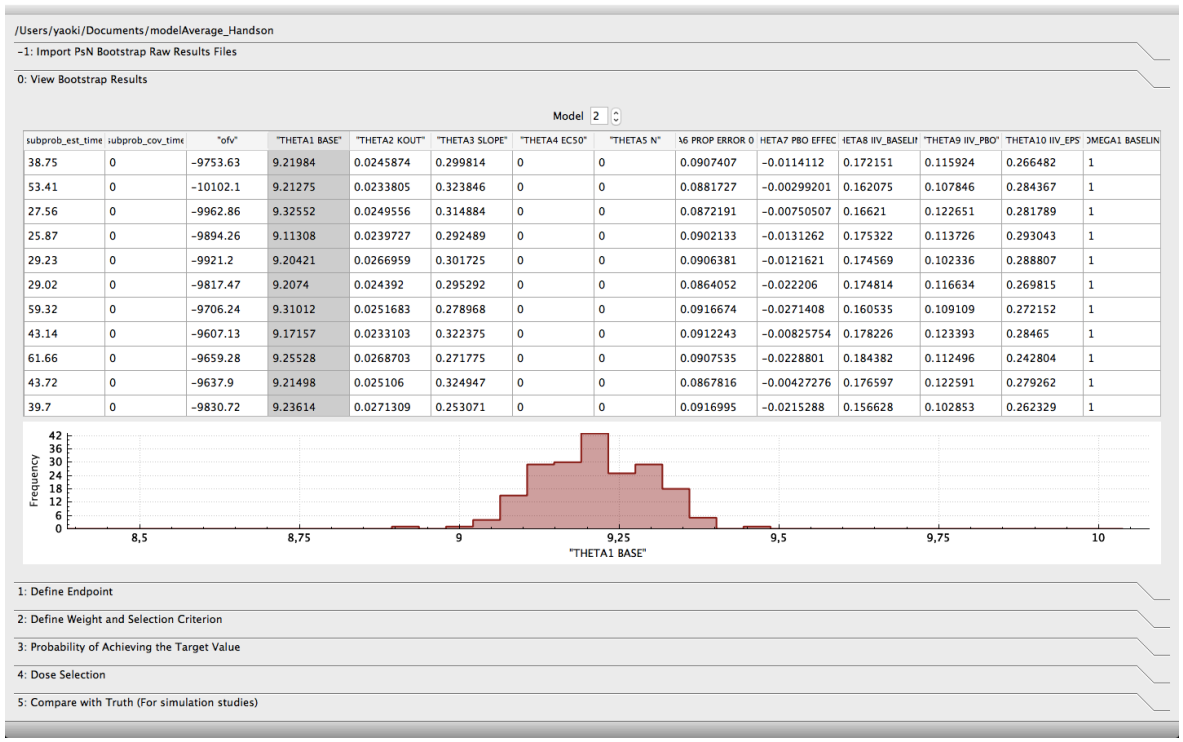
We now assume that all the bootstrap runs have successfully completed using the shell-script created at exercise 1. We have provided the following raw_results files generated by PsN bootstrap.

The raw_result files were generated based on a simulated PhIIb study data of Plasma Glucose measurement in Patients with Type2 Diabetes following a treatment (cf. http://wwwdev.ebi.ac.uk/biomodels/model-repository/model/DDMODEL00000079). The dose-response models used for the bootstrap analyses are described in detail in Section 5 of this document.

Import the raw_result files that are provided with this hands-on.



3: Visualize Parameter Distributions



exercise 1

For each model make observations on the distribution of the model parameters. Discuss if bootstrap was really necessary and if we can approximate the parameter distribution reasonably well with the variance-covariance matrix.

(hint)

Take a look at the parameter distributions of Model 3 and 4 (emax and sigmoidal).

4: Endpoint

One key assumption we made in this software is that the end-point (or the primary efficacy variable) can be defined using the values appear in a raw_results file, mainly fixed effect parameters (THETAs) and distribution parameters of the random effect variables (ETAs).

In order to define the endpoint this way, we may need to do a small mathematical manipulation to the original nonlinear mixed effect models. We consider the following nonlinear mixed effect models in this hands on exercise.

Baseline (Initial Condition)

$$A|_{t=0} = \theta_1 e^{\theta_8 \eta_1}$$

Model 0: Placebo

$$\frac{dA}{dt} = \theta_2 \theta_1 e^{\theta_8 \eta_1}(1 + \theta_7 + \theta_9 \eta_2) - \theta_2 A\,(1 + 0)$$

Model 1: Linear

$$\frac{dA}{dt} = \theta_2 \theta_1 e^{\theta_8 \eta_1}(1 + \theta_7 + \theta_9 \eta_2) - \theta_2 A(1 + \theta_3 \mathrm{dose})$$

Model 2: Log Linear

$$\frac{dA}{dt} = \theta_2 \theta_1 e^{\theta_8 \eta_1}(1 + \theta_7 + \theta_9 \eta_2) - \theta_2 A(1 + \theta_3 \ln(\mathrm{dose} + 1))$$

Model 3: Emax

$$\frac{dA}{dt} = \theta_2 \theta_1 e^{\theta_8 \eta_1}(1 + \theta_7 + \theta_9 \eta_2) - \theta_2 A\left(1 + \theta_3 \frac{\mathrm{dose}}{\theta_4 + \mathrm{dose}}\right)$$

Model 4: Sigmoidal

$$\frac{dA}{dt} = \theta_2 \theta_1 e^{\theta_8 \eta_1}(1 + \theta_7 + \theta_9 \eta_2) - \theta_2 A\left(1 + \theta_3 \frac{\mathrm{dose}^{\theta_5}}{\theta_4^{\theta_5} + \mathrm{dose}^{\theta_5}}\right)$$

The end-point of this study is the baseline-adjusted steady state value of A (expectation value of the difference between the steady state value of A and the initial state of A).

exercise 2
Derive a closed form expression (in terms is theta, eta, and dose) of the endpoint for each model.

(bonus question)
Assuming all etas follow normal distribution of mean 0 variance 1, discuss if the population mean/median/75percentile of the endpoint can be expressed analytically. If it cannot be expressed analytically, propose a way to approximate these quantities numerically.

(hint)
Assuming the steady state exists, you can let the right hand side of the ODE equal to zero to derive the steady state of the ODE, i.e.,

Assume

$$\frac{dA_{\text{steady}}}{dt} = 0$$

Hence (for example for Model 1)

$$0 = \theta_2\theta_2 e^{\theta_8\eta_1}(1 + \theta_7 + \theta_9\eta_2) - \theta_2 A_{\text{steady}}(1 + \theta_3\text{dose})$$

Rewrite the above equation into the form

$$A_{\text{steady}} = \dots$$

Then obtain the baseline adjusted value as in

$$\text{End Point} := (A_{\text{steady}} - A|_{t=0}) = \dots$$

(hint)
The model1 is defined in the NONMEM model files as follows:

```
BASE=THETA(1)*EXP(THETA(8)*ETA(1))

PBO    = THETA(7) +THETA(9)*ETA(2)

DREFF  = THETA(3)*DOSE

KIN = THETA(2)*BASE

F1=BASE

$DES

DADT(1)=KIN*(1+PBO)-KOUT*A(1)*(1+DREFF)
```

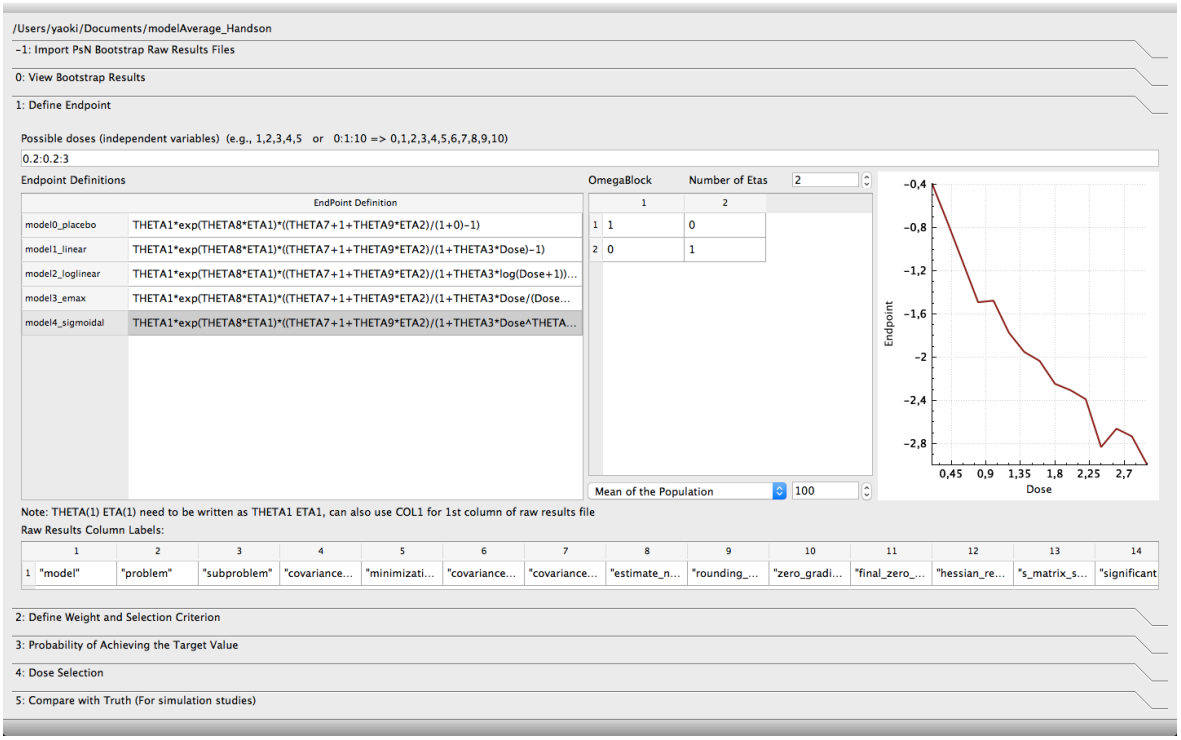
(answer...)

Model 0: THETA1*exp(THETA8*ETA1)*((THETA7+1+THETA9*ETA2)/(1+0)-1)
Model 1: THETA1*exp(THETA8*ETA1)*((THETA7+1+THETA9*ETA2)/(1+THETA3*Dose)-1)
Model 2: THETA1*exp(THETA8*ETA1)*((THETA7+1+THETA9*ETA2)/(1+THETA3*log(Dose+1))-1)
Model 3: THETA1*exp(THETA8*ETA1)*((THETA7+1+THETA9*ETA2)/(1+THETA3*Dose/(Dose
+THETA4))-1)
Model 4: THETA1*exp(THETA8*ETA1)*((THETA7+1+THETA9*ETA2)/(1+THETA3*Dose^THETA5/
(Dose^THETA5+THETA4^THETA5))-1)

5: Define Endpoint

Input the endpoint we have derived in 4 into “modelAverage”. Keep in mind that THETA(1) need to be written as THETA1 and ETA(1) need to be written as ETA1. We assume the random effect variables in the model (ETAs) follow the normal distribution of mean 0 variance 1 (hence we do not need to modify the Omega matrix).

The does of interest of this dose finding study is between 0 and 3 with the accuracy of 0.2 dose unit.



exercise 3

Make some observations on the relationship between the stochastic noise and the number of simulations. Think of how this stochastic noise would influence the accuracy of the dose selection.

Given that you would like to choose the dose within 0.2 dose unit of accuracy, how many simulations do you need?

6: Define Weights

In modelAverage, the weight for the average and model selection criteria can be defined flexibly. For example using Akaike Information Criteria can be used assuming the sample size is large enough. AIC is defined as follows:

AIC= 2 x number of the fixed effect parameters - 2 ln(Likelihood)

exercise 4
Input the weight definition of your choice.
Please keep in mind that the negative number, for example -1, needs to be written as (0-1) in the current implementation of the software (sorry!).

Discuss as to why we need to define weight in log scale. (i.e., one can experiment defining the weight in linear scale by unchecking “in log scale” checkbox and see what happens.)

(hint)
Normalized Weight using AIC for model1 can be defined as follows:

$$\text{Normalized Weight}_{\text{model1}} = \frac{\exp\left(\frac{\text{AIC}_{\text{model1}}}{-2}\right)}{\sum_{i=0}^{4}\exp\left(\frac{\text{AIC}_{\text{model}i}}{-2}\right)}$$

By default the modelAverage takes the weight that is defined in natural log scale.

(bonus question)
What is the smallest and the largest value “double number” can take?
Note that both the numerator and denominator of the Normalized Weights defined above are larger than the largest value “double number” can take.
Suggest a way to compute the weight using double precision computation.

/Users/yaoki/Documents/modelAverage_Handson

-1: Import PsN Bootstrap Raw Results Files

0: View Bootstrap Results

1: Define Endpoint

2: Define Weight and Selection Criterion

Weight expression (assumes ofv is at 20th column, leave blank to exclude the model) ☑ in log scale

	Weight Definition in log scale
model0_placebo	(ofv+0)/(0-2)
model1_linear	(ofv+2)/(0-2)
model2_loglinear	(ofv+2)/(0-2)
model3_emax	(ofv+4)/(0-2)
model4_sigmoidal	(ofv+6)/(0-2)

Relative weight

	model0_placebo	model1_linear	model2_loglinear	model3_emax	model4_sigmoida
1	2.61348e-77	0.316755	0.0997088	0.412692	0.170844
2	1.02462e-81	0.54941	0.0148203	0.292098	0.143672
3	3.75444e-87	0.459451	0.0405881	0.361205	0.138755
4	4.05529e-72	0.539819	0.0257539	0.31174	0.122687
5	9.61023e-87	0.305135	0.151493	0.389274	0.154098
6	8.73824e-79	0.592775	0.0174443	0.281039	0.108742
7	7.22762e-69	0.580501	0.0268703	0.279099	0.113531
8	9.29944e-82	0.488208	5.90094e-06	0.1796	0.332186

3: Probability of Achieving the Target Value

4: Dose Selection

5: Compare with Truth (For simulation studies)

7: Obtain the Dose v.s. PoS graph

We are now at the last stage of this Hands on... You can simply choose the way you want to do the model averaging, and define the target endpoint (you may use inf or -inf as one of the boundaries for the target value).

exercise 5
Obtain the probability of achieving the target endpoint value v.s. dose graph. In this exercise, the target effect is defined as follows:

Baseline adjusted decrease of more than 2 units in FPG at the steady state.

Observe if 200 bootstrap sampling is enough or need more bootstrap samples to reliably calculate the PoS.
Make your dose selection using this graph.