Variable selection / regression coefficients estimation:

Methods:

Exactly same methods as with two classes of bmi, but having three classes of bmi:
0 bmi<25,
1 25>=bmi<30
2 bmi>=30

modeling with multinom from nnet package in R
and calculating the AUC for multiclass, by taking the average of three AUC for each class.

Pseudo R^2 calculated as:
1 - ration of likelihood of model and likelihood of null model(McFadden's R^2)

AUC difference significance calculated for each class pair(3x).

Results:

Visit 1 basic covariates (age, agesq, gender_factor, year, ffq_factor):
AUC 0.6206, AIC 54,872, R^2 0.02863

Visit 1 itself:

	diet beta1	diet beta2	p-value	AIC	AUC	R^2
alpha 0	1.856	3.750	< 2e-16	54,024	0.6467	0.04372
alpha 0.1	1.845	3.747	< 2e-16	54,013	0.6470	0.04392
alpha 0.2	1.807	3.675	< 2e-16	54,009	0.6471	0.04399
alpha 0.3	1.807	3.673	< 2e-16	54,010	0.6471	0.04397
alpha 0.4	1.801	3.664	< 2e-16	54,010	0.6471	0.04397
alpha 0.5	1.801	3.668	< 2e-16	54,010	0.6471	0.04396
alpha 0.6	1.762	3.588	< 2e-16	54,019	0.6469	0.04382
alpha 0.7	1.769	3.602	< 2e-16	54,018	0.6469	0.04382
alpha 0.8	1.777	3.622	< 2e-16	54,018	0.6469	0.04383
alpha 0.9	1.771	3.610	< 2e-16	54,018	0.6469	0.04383
alpha 1	1.767	3.604	< 2e-16	54,015	0.6470	0.04389
fitting separately all	0.273	0.556	< 2e-16	54,431	0.6354	0.03652
fitting together all	1.263	2.530	< 2e-16	53,955	0.6488	0.04495

fitting separately selected	0.414	0.851	< 2e-16	54,385	0.6369	0.03734
fitting together selected	1.337	2.723	< 2e-16	54,011	0.6476	0.04395

Selected variables in best model, same as before:

MONOsum1, mfetsum1, FA, protsum1_anim, protsum1_veg, DISAsum1, TRANSsum1,NATRsum1 ensum1, MAGNsum1, FOSFsum1,ZINCsum1, retisum1,karosum1, TIAMsum1,Folasum1, B6sum1, B12sum1, askosum1,Dsum1, tokosum1,VITKsum1,JODIsum1, KALIsum1

Visit 2:

Visit 2 basic covariates (age, agesq, gender_factor, year, ffq_factor):
AUC 0.6040, AIC 61,884, R^2 0.01976

predicting with models from visit1, using basic covariates and diet score from visit 2:

	AUC
alpha 0	0.6191
alpha 0.1	0.6212
alpha 0.2	0.6222
alpha 0.3	0.6224
alpha 0.4	0.6227
alpha 0.5	0.6228
alpha 0.6	0.6221
alpha 0.7	0.6223
alpha 0.8	0.6225
alpha 0.9	0.6226
alpha 1	0.6229
fitting separately	0.6004
fitting together	0.6274

modeling obesity in visit 2 with diet score from visit1:

	diet beta1	diet beta2	p-value	AIC	AUC	R^2
alpha 0	1.642	3.418	< 2e-16	61,036	0.6343	0.03326
alpha 0.1	1.608	3.382	< 2e-16	61,038	0.6342	0.03324
alpha 0.2	1.573	3.313	< 2e-16	61,037	0.6342	0.03326
alpha 0.3	1.573	3.316	< 2e-16	61,037	0.6342	0.03324
alpha 0.4	1.569	3.306	< 2e-16	61,037	0.6343	0.03325
alpha 0.5	1.571	3.309	< 2e-16	61,038	0.6343	0.03324
alpha 0.6	1.537	3.246	< 2e-16	61,045	0.6341	0.03312
alpha 0.7	1.543	3.258	< 2e-16	61,045	0.6341	0.03312
alpha 0.8	1.552	3.274	< 2e-16	61,044	0.6342	0.03314
alpha 0.9	1.547	3.264	< 2e-16	61,044	0.6342	0.03313
alpha 1	1.544	3.257	< 2e-16	61,042	0.6342	0.03317
fitting separately	0.361	0.782	< 2e-16	61,390	0.6243	0.02765
fitting together	1.128	2.436	< 2e-16	61,052	0.6334	0.03301

modeling obesity in visit 2 with basic covariates and diet score from visit2:

	diet beta1	diet beta2	p-value	AIC	AUC	R^2
alpha 0.3	1.651	3.617	< 2e-16	60,757	0.6394	0.03769
fitting separately	0.350	0.783	< 2e-16	61,279	0.6259	0.02942
fitting together	1.358	3.005	< 2e-16	60,656	0.6417	0.03929

Prediction tables for each visit, by taking regression coefficients from the independent dataset and constructing the diet score in that visit. Green=”right”, Red=”wrong” = “our interest”

Shrinkage model, alpha 0.3:

visit 1:

true / predicted	0	1	2
0	14501	7514	1910
1	2825	3656	774
2	2	0	1

visit 2:

true / predicted	0	1	2
0	8251	5477	2143
1	5002	7694	2501
2	37	44	34

OLS model, all significant variables:

visit 1:

true / predicted	0	1	2
0	14545	7560	1819
1	2761	3585	849
2	22	25	17

visit 2:

true / predicted	0	1	2
0	6347	3984	1406
1	6752	9005	3114
2	191	226	158