BASIC ASSIGNMENT 4
CORE METHODS IN EDUCATIONAL DATA MINING
PROFESSOR BAKER
PERFORMANCE FACTOR ANALYSIS
DUE NOON, MONDAY OCTOBER 27

In this assignment, you need to build a Performance Factors Analysis model for data file pfa-modelfit-set-v3.xlsx

This data set's variables are:

- Student – a deidentified ID for the student
- Item – the problem step in the learning system
- Skill1,2,3 – does this problem step involve the current skill (1) or not (0), based upon the Q-Matrix
- Firstattempt-correctness – did the student get the item right on the first attempt?
- Firstattempt-incorrect – did the student get the item wrong on the first attempt?

You should complete questions 1-10 of this assignment in Microsoft Excel, or a similar spreadsheet program.

Question 1:

The first thing we need to do is to create a column that represents the success so far on skill 1, 2, and 3. This will be used with PFA's gamma parameter. We'll put these in columns H, I, and J. What should go in cell H2? (Remember, if you're not sure, try each of these)

A) =IF($C2=1,$F2,0),IF($C2=1,H1+$F2,H1)
B) =IF($A2<>$A1,IF($C2=1,$F2,0),IF($C2=1,H1+$F2,H1))
C) =IF($C2=0,$F2,1),IF($C2=1,H1+$F2,H1)
D) =IF($A2<>$A1,IF($C2=0,$F2,1),IF($C2=1,H1+$F2,H1))
E) =IF($C2=0,$F2,1),IF($C2=1,H1, H1+$F2)
F) =IF($A2<>$A1,IF($C2=0,$F2,1),IF($C2=1,H1, H1+$F2))
G) =IF($C2=1,$F2,0),IF($C2=1,H1*$F2,H1)
H) =IF($A2<>$A1,IF($C2=1,$F2,0),IF($C2=1,H1*$F2,H1))

Question 2:

Copy H2 down and make the corresponding version of columns I and J. Next, you need to create a column that represents the incorrect answers so far on skill 1, 2, and 3. This will be used with PFA's rho

parameter. We'll put these in columns K, L, and M. What should go in cell K2? (Remember, if you're not sure, try each of these)

A) =IF($A2<>$A3,IF(C2=1,$G2,0),IF(C2=1,K1+$G2,1))
B) =IF($A2<>$A3,IF(C2=1,$G2,0),IF(C2=1,K1+$G2,0))
C) =IF($A2<>$A3,IF(C2=1,$G2,0),IF(C2=1,K1+$G2,K1-$G2))
D) =IF($A2<>$A3,IF(C2=1,$G2,0),IF(C2=1,K1+$G2,K1))
E) =IF(C2=1,$G3,0)
F) =IF($A2<>$A1,IF(C2=1,$G2,0),IF(C2=1,K1+$G2,K1))
G) =IF($A2<>$A1,IF(C2=1,$G2,0),IF(C2=1,K1,K1+$G2))
H) =IF($A2<>$A3,IF(C2=1,$G2,0),IF(C2=1,K1,K1+$G2))

Question 3. Now you need to compute the gamma parameters for the student's history of success. Note that the gamma weights are on sheet "fit". Copy =fit!$F$1*H2 into cell N2 and propagate it down. What should O2 be?

A) =fit!$F$2*I2
B) =fit!$F$1*H2
C) =fit!$F$1*I2
D) =fit!$F$2*H2

Question 4. OK, propagate from O2 down, and do the same thing for column P. Put =SUM(N2:P2) into cell Q2 and copy it down. Now you have all the success parameters added together for the three skills. Now you need to create the rho parameters for the student's history of failure. What should R2 be?

A) =fit!$F$2*K2
B) =fit!$F$2*L2
C) =fit!$F$2*M2
D) =fit!$F$3*L2
E) =fit!$F$4*K2
F) =fit!$F$5*L2
G) =fit!$F$6*M2

Question 5. Propagate that down, and do the same thing for column S and T. Put =SUM(R2:T2) into cell U2 and copy it down. Now you have all the failure parameters added together for the three skills. Put =fit!$F$7 into column V and copy it down. Now we can calculate m! (If you don't remember what m is, re-watch the video). What do you put in cell W2?

A) =V2
B) =Q2+U2
C) =Q2/U2
D) =Q2*U2
E) =Q2+U2+V2
F) =Q2*U2*V2

G) =(Q2*U2)/V2
H) =EXP(Q2+U2-V2)
I) =EXP(Q2+U2+V2)
J) =EXP(Q2*U2*V2)

Question 6. Propagate that down. What goes in X2?

A) =W2
B) =(W2*-1)
C) =EXP(W2)
D) =EXP(W2*-1)
E) =(1+EXP(W2))
F) =(1+EXP(W2*-1))
G) =(1-EXP(W2))
H) =(1-EXP(W2*-1))
I) =1/(1-EXP(W2))
J) =1/(1+EXP(W2))
K) =1/(1-EXP(W2*-1))
L) =1/(1+EXP(W2*-1))

Question 7. Propagate that down. You've got PFA! Now it's time to fit the seven parameters. Go to the sheet "fit". What is the SSR currently?

Correct answer: 255. Yes, this is here intentionally.

Question 8. What happens if you change gamma-skill 1 to be 1? What does the SSR become?

Question 9. Is the model better or worse than the model you got for question 7?

A) Better
B) Worse
C) The Same

Question 10. What does it mean to increase gamma-skill-1 from 0 to 1?

Question 11. How good can you get the model (best SSR) by changing any parameter to -1, 0, or 1?

Question 12. Use the Excel Equation Solver to find the optimal parameters for this model. (You may need to install it as an add-in). Make sure to use the GRG Nonlinear solving method and leave make unconstrained variables non-negative unchecked. What is the resultant SSR?