Learning curves for traffic controllers

Contents

1. Introduction

Kanfer & Ackerman (1989) reported on data for 141 U.S. Air Force enlisted personnel who carried out a computerized air traffic controller task developed by the authors. Please note that SSI is using the data with permission of the authors. The data remain the copyrighted property of the Kanfer and Ackerman, and further publication or further dissemination of these data is not permitted without the expressed consent of the copyright owners.

The subjects performed air traffic controlling tasks: accepting planes into their hold pattern and landing them safely and efficiently on one of four runways varying in length and compass directions, subject to rules governing plane movements and landing requirements. In the data set shown below, the success of a series of between three and six ten minute trials are recorded. The measure of success was the number of correct landings in a trial.

trial.lsf

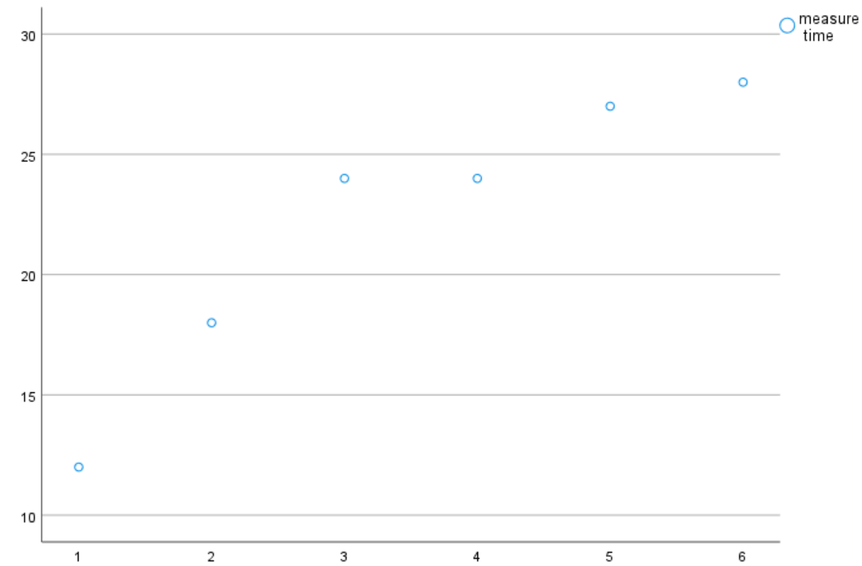
	TRIAL1	TRIAL2	TRIAL3	TRIAL4	TRIAL5	TRIAL6	ABILITY
1	24.00	27.00	30.00	32.00	38.00	41.00	1.42
2	2.00	3.00	9.00	13.00	13.00	14.00	-0.08
3	12.00	18.00	24.00	24.00	27.00	28.00	-0.67
4	7.00	19.00	23.00	30.00	30.00	32.00	0.11
5	17.00	24.00	28.00	29.00	33.00	34.00	-0.02
6	14.00	23.00	27.00	26.00	29.00	28.00	2.32
7	18.00	24.00	29.00	31.00	34.00	37.00	0.92
8	9.00	14.00	18.00	22.00	25.00	24.00	-1.22
9	3.00	4.00	27.00	22.00	27.00	24.00	-1.11
10	6.00	27.00	25.00	34.00	39.00	44.00	0.14

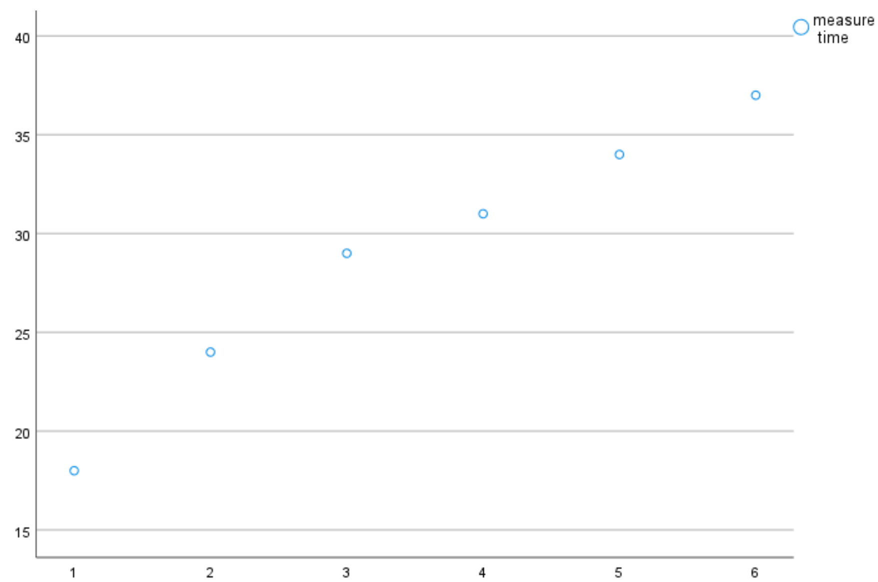
We note a lot of variation in the number of successes in the initial trials that decrease towards to late trials. Also note that some controllers only completed a subset of trials, and that the missing values are coded as -9. While it is possible to deal with the missing data using FIML, the small number of missing values (7 out of 987) we opt to simply replace the missing values by the means of the variables. After this replacement, we now have the data set **trial_filled.lsf** and that is the data we will use in this example.

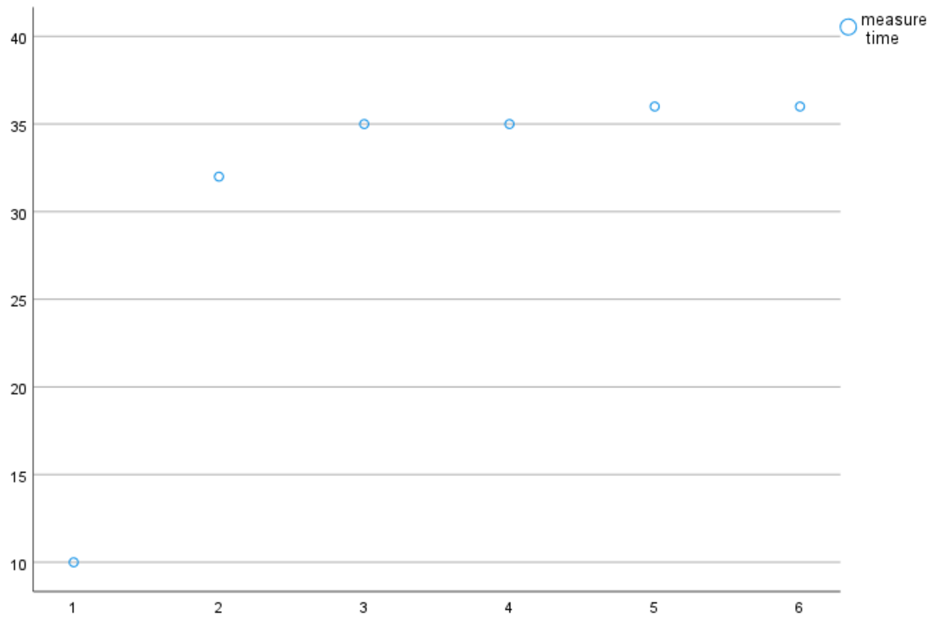
Additional information for each respondent is available in the form of the results from the Armed Services Vocational Battery test. This score is included as the variable ABILITY in the data set.

2. Quadratic growth

Below are scatterplots for the observed outcome over time for three of the respondents. From these, it looks as if the growth curves may not necessarily be linear.







We first fit a model with a quadratic term to these data. The model can be specified as

$$y_{it} = a_i + b_{i1}(TIME)_{it} + b_{i2}(TIME)_{it}^2 + e_{it}$$

where $Time_t = t-1, t = 1, 2, ..., 6.$

In LISREL notation, this model is written as

$$\mathbf{x} = \mathbf{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}$$

with $\mathbf{x} = (y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5}, y_{i6})$ and $\boldsymbol{\xi} = (a_i, b_{i1}, b_{i2})$ and where

$$\mathbf{\Lambda}_x = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ \vdots & \vdots & \vdots \\ 1 & 5 & 25 \end{bmatrix}.$$

$Cov(\boldsymbol{\delta}) = \sigma_e^2 \mathbf{I} = Var(e_{it})\mathbf{I}.$ The mean vector and covariance matrix of x is

$$\boldsymbol{\mu} = \mathbf{\Lambda}_x \boldsymbol{\kappa},$$
$$\mathbf{\Sigma} = \mathbf{\Lambda}_x \mathbf{\Phi} \mathbf{\Lambda}_x' + \sigma_e^2 \mathbf{I}.$$

The parameter matrices are $\boldsymbol{\kappa}$, $\mathbf{\Phi}$ and σ_e^2.

The following SIMPLIS syntax may be used to fit a quadratic growth curve to the data:

```
Quadratic Growth Curve for TRial Data
Raw Data from File TRIAL_filled.LSF
Missing Value Code -9
Latent Variables: a b1 b2
Relationships
TRIAL1 = 1*a 0*b1   0*b2
TRIAL2 = 1*a 1*b1   1*b2
TRIAL3 = 1*a 2*b1   4*b2
TRIAL4 = 1*a 3*b1   9*b2
TRIAL5 = 1*a 4*b1  16*b2
TRIAL6 = 1*a 5*b1  25*b2
a b1 b2= CONST
Equal Error Variances: TRIAL1 - TRIAL6
Path Diagram
End of Problem
```

The sample means, variance and covariances are reported in the output file.

Covariance Matrix

	TRIAL1	TRIAL2	TRIAL3	TRIAL4	TRIAL5	TRIAL6
TRIAL1	53.249					
TRIAL2	49.079	76.090				
TRIAL3	39.356	63.388	79.913			
TRIAL4	33.649	58.070	72.854	85.335		
TRIAL5	31.581	54.008	66.862	76.019	81.839	
TRIAL6	32.392	56.316	68.276	78.350	80.401	94.133

Means

TRIAL1	TRIAL2	TRIAL3	TRIAL4	TRIAL5	TRIAL6
11.631	21.156	27.305	30.878	32.392	34.196

The mean vector and covariance matrix of independent variables indicate that our suspicion that a linear curve may not be appropriate is correct.

Mean Vector of Independent Variables

a	b1	b2
12.137	9.391	-1.020
(0.645)	(0.390)	(0.062)
18.822	24.055	-16.451

Covariance Matrix of Independent Variables

	a	b1	b2
a	50.464		
	(7.004)		
	7.205		

```
b1      -4.003     14.270
        (3.128)    (2.607)
        -1.280      5.474

b2       0.053     -1.846      0.276
        (0.485)    (0.403)    (0.067)
         0.110    -4.578      4.116
```

3. Cubic growth curve

As a next step, we fit a cubic growth curve to the data. This implies extending the model to

$$y_{it} = a_i + b_{i1}(TIME)_{it} + b_{i2}(TIME)_{it}^2 + b_{i3}(TIME)_{it}^3 + e_{it}$$

This implies

$$\mathbf{\Lambda}_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 5 & 25 & 125 \end{bmatrix}$$

as reflected in the syntax file **trial1a.spl**:

trial1a.spl

```
Cubic Growth Curve for TRIAL Data
Raw Data from File TRIAL_filled.LSF
Latent Variables: a b1 b2 b3
Relationships
TRIAL1 = 1*a 0*b1   0*b2    0*b3
TRIAL2 = 1*a 1*b1   1*b2    1*b3
TRIAL3 = 1*a 2*b1   4*b2    8*b3
TRIAL4 = 1*a 3*b1   9*b2   27*b3
TRIAL5 = 1*a 4*b1  16*b2   64*b3
TRIAL6 = 1*a 5*b1  25*b2  125*b3
a b1 b2 b3= CONST
Equal Error Variances: TRIAL1 - TRIAL6
Path Diagram
End of Problem
```

The output for this model is as follows.

```
Covariance Matrix of Independent Variables

               a         b1         b2         b3
        --------   --------   --------   --------
  a       46.766
         (6.516)
           7.177

b1         5.349     30.196
         (4.864)    (7.455)
           1.100      4.050
```

```
b2     -5.607     -9.650      4.238
       (2.125)    (3.162)    (1.484)
       -2.639     -3.052      2.856

b3      0.798      0.913     -0.471      0.056
       (0.267)    (0.379)    (0.185)    (0.024)
        2.985      2.411     -2.539      2.358


  Mean Vector of Independent Variables

            a         b1         b2         b3
     --------   --------   --------   --------
       11.584     11.913     -2.401      0.184
      (0.621)    (0.648)    (0.284)    (0.036)
       18.643     18.392     -8.448      5.162
```

Since the estimated means are all significant as are the variances, we conclude that there is considerable variation in the shape of the growth curves over respondents.

Looking over the measurement equations, we note a monotone increase in R^2 over the successive trials. The error variance is estimated at 7.992 and is highly significant.

```
  TRIAL1 = 1.000*a, Errorvar.= 7.992 , R² = 0.854
Standerr                      (0.675)
Z-values                       11.832
P-values                       0.000

  TRIAL2 = 1.000*a + 1.000*b1 + 1.000*b2 + 1.000*b3, Errorvar.= 7.992 , R² = 0.889
Standerr                                                        (0.675)
Z-values                                                         11.832
P-values                                                         0.000

  TRIAL3 = 1.000*a + 2.000*b1 + 4.000*b2 + 8.000*b3, Errorvar.= 7.992 , R² = 0.901
Standerr                                                        (0.675)
Z-values                                                         11.832
P-values                                                         0.000

  TRIAL4 = 1.000*a + 3.000*b1 + 9.000*b2 + 27.000*b3, Errorvar.= 7.992 , R² = 0.904
Standerr                                                         (0.675)
Z-values                                                          11.832
P-values                                                          0.000

  TRIAL5 = 1.000*a + 4.000*b1 + 16.000*b2 + 64.000*b3, Errorvar.= 7.992 , R² = 0.907
Standerr                                                          (0.675)
Z-values                                                           11.832
P-values                                                           0.000

  TRIAL6 = 1.000*a + 5.000*b1 + 25.000*b2 + 125.000*b3, Errorvar.= 7.992 , R² = 0.915
Standerr                                                           (0.675)
Z-values                                                            11.832
P-values                                                            0.000
```

We conclude that the fit may be improve by allowing the error variances to be different. To do so, we use the syntax file. note the removal of the equality constraint on the error variances of TRIAL1 to TRIAL6.

trial1aa.spl

```
Cubic Growth Curve for TRIAL Data
Raw Data from File TRIAL_filled.LSF
Latent Variables: a b1 b2 b3
Relationships
TRIAL1 = 1*a 0*b1   0*b2    0*b3
TRIAL2 = 1*a 1*b1   1*b2    1*b3
TRIAL3 = 1*a 2*b1   4*b2    8*b3
TRIAL4 = 1*a 3*b1   9*b2   27*b3
TRIAL5 = 1*a 4*b1  16*b2   64*b3
TRIAL6 = 1*a 5*b1  25*b2  125*b3
a b1 b2 b3= CONST
Path Diagram
End of Problem
```

Results are as follows:

```
  TRIAL1 = 1.000*a, Errorvar.= -12.597, R² = 1.235
Standerr                      (7.804)
Z-values                      -1.614
P-values                       0.107

W_A_R_N_I_N_G : Error variance is negative.

  TRIAL2 = 1.000*a + 1.000*b1 + 1.000*b2 + 1.000*b3, Errorvar.= 15.310, R² = 0.804
Standerr                                                      (2.643)
Z-values                                                       5.792
P-values                                                       0.000

  TRIAL3 = 1.000*a + 2.000*b1 + 4.000*b2 + 8.000*b3, Errorvar.= 7.205 , R² = 0.910
Standerr                                                      (1.643)
Z-values                                                       4.385
P-values                                                       0.000

  TRIAL4 = 1.000*a + 3.000*b1 + 9.000*b2 + 27.000*b3, Errorvar.= 7.209 , R² = 0.913
Standerr                                                       (1.428)
Z-values                                                        5.050
P-values                                                        0.000

  TRIAL5 = 1.000*a + 4.000*b1 + 16.000*b2 + 64.000*b3, Errorvar.= 5.170 , R² = 0.938
Standerr                                                        (1.681)
Z-values                                                         3.075
P-values                                                         0.002

  TRIAL6 = 1.000*a + 5.000*b1 + 25.000*b2 + 125.000*b3, Errorvar.= 4.795 , R² = 0.949
Standerr                                                         (6.263)
Z-values                                                          0.766
P-values                                                          0.444
```

We note that the error variance associated with the first trial variable is negative, though not significant.

```
        Covariance Matrix of Independent Variables

                   a          b1          b2          b3
            --------    --------    --------    --------
       a      66.136
            (10.669)
               6.199

      b1     -19.578      57.633
            (10.756)    (12.820)
              -1.820       4.496

      b2       3.537     -18.942       7.286
             (3.798)     (4.600)     (1.809)
               0.931      -4.118       4.028

      b3      -0.194       1.882      -0.787       0.089
             (0.412)     (0.500)     (0.211)     (0.026)
              -0.470       3.761      -3.727       3.387


         Mean Vector of Independent Variables

                   a          b1          b2          b3
            --------    --------    --------    --------
              11.706      11.926      -2.444       0.191
             (0.612)     (0.648)     (0.282)     (0.035)
              19.121      18.405      -8.675       5.468


                          Goodness-of-Fit Statistics

 Degrees of Freedom for (C1)-(C2)                             7
 Maximum Likelihood Ratio Chi-Square (C1)                     7.486 (P = 0.3800)
 Due to Covariance Structure                                  5.377
 Due to Mean Structure                                        2.109
 Browne's (1984) ADF Chi-Square (C2_NT)                       7.377 (P = 0.3907)
```

The goodness-of-fit measures indicate an acceptable fit.

To deal with the negative error variance for the first trial in the previous model, we set this error variance to a small value. In the syntax shown below, we have set the error variance for TRIAL1 to 1.

```
Cubic Growth Curve for TRIAL Data
Raw Data from File TRIAL_filled.LSF
Latent Variables: a b1 b2 b3
Relationships
TRIAL1 = 1*a 0*b1   0*b2    0*b3
TRIAL2 = 1*a 1*b1   1*b2    1*b3
TRIAL3 = 1*a 2*b1   4*b2    8*b3
TRIAL4 = 1*a 3*b1   9*b2   27*b3
TRIAL5 = 1*a 4*b1  16*b2   64*b3
TRIAL6 = 1*a 5*b1  25*b2  125*b3
a b1 b2 b3= CONST
Set the Error Variance of TRIAL1 to 1
Path Diagram
End of Problem
```

The measurement equations and goodness-of-fit measures for this model are given below.

```
  TRIAL1 = 1.000*a, Errorvar.= 1.000, R² = 0.981

  TRIAL2 = 1.000*a + 1.000*b1 + 1.000*b2 + 1.000*b3, Errorvar.= 12.944, R² = 0.828
Standerr                                                      (2.089)
Z-values                                                       6.198
P-values                                                       0.000

  TRIAL3 = 1.000*a + 2.000*b1 + 4.000*b2 + 8.000*b3, Errorvar.= 7.997 , R² = 0.901
Standerr                                                      (1.597)
Z-values                                                       5.008
P-values                                                       0.000

  TRIAL4 = 1.000*a + 3.000*b1 + 9.000*b2 + 27.000*b3, Errorvar.= 7.018 , R² = 0.915
Standerr                                                       (1.433)
Z-values                                                        4.899
P-values                                                        0.000

  TRIAL5 = 1.000*a + 4.000*b1 + 16.000*b2 + 64.000*b3, Errorvar.= 5.558 , R² = 0.934
Standerr                                                        (1.695)
Z-values                                                         3.279
P-values                                                         0.001

  TRIAL6 = 1.000*a + 5.000*b1 + 25.000*b2 + 125.000*b3, Errorvar.= 3.943 , R² = 0.958
Standerr                                                         (6.252)
Z-values                                                          0.631
P-values                                                          0.528
                         Goodness-of-Fit Statistics

Degrees of Freedom for (C1)-(C2)                     8
Maximum Likelihood Ratio Chi-Square (C1)             10.805 (P = 0.2130)
Due to Covariance Structure                          8.707
Due to Mean Structure                                2.098
Browne's (1984) ADF Chi-Square (C2_NT)               10.994 (P = 0.2020)
```

We note that the model still provides an acceptable fit.

4. Cubic growth with covariate

Additional information on the respondents is available in their scores on the Armed Services Vocational Battery test. To investigate whether this measure have an impact on the shape of the growth curve, we now add the variable ABILITY as covariate.

trial2a.spl

```
Cubic Growth Curve with Covariate for TRIAL Data
Raw Data from File TRIAL_filled.LSF
Latent Variables: a b1 b2 b3
Relationships
TRIAL1 = 1*a 0*b1   0*b2    0*b3
TRIAL2 = 1*a 1*b1   1*b2    1*b3
TRIAL3 = 1*a 2*b1   4*b2    8*b3
TRIAL4 = 1*a 3*b1   9*b2   27*b3
TRIAL5 = 1*a 4*b1  16*b2   64*b3
TRIAL6 = 1*a 5*b1  25*b2  125*b3
a b1 b2 b3= CONST ABILITY
Let the errors of a-b3 correlate
Set the Error Variance of TRIAL1 to 1
Path Diagram
End of Problem
```

For this model, the following structural equation results are obtained.

```
Structural Equations


        a = 11.625 + 3.686*ABILITY, Errorvar.= 38.764, R² = 0.259
 Standerr  (0.531)  (0.531)                   (4.735)
 Z-values   21.892   6.941                     8.186
 P-values   0.000    0.000                     0.000

       b1 = 11.982 + 0.624*ABILITY, Errorvar.= 39.047, R² = 0.00988
 Standerr  (0.649)  (0.649)                   (7.346)
 Z-values   18.470   0.962                     5.315
 P-values   0.000    0.336                     0.000

       b2 =  - 2.453  - 0.554*ABILITY, Errorvar.= 4.925 , R² = 0.0587
 Standerr     (0.278)  (0.278)                   (1.436)
 Z-values     -8.827   -1.994                     3.430
 P-values      0.000    0.046                     0.001

       b3 = 0.192   + 0.0784*ABILITY, Errorvar.= 0.0632 , R² = 0.0887
 Standerr  (0.0344)  (0.0344)                   (0.0241)
 Z-values   5.577     2.282                      2.622
 P-values   0.000     0.022                      0.009
```

While 26% of the variation in the intercept is explained by the current model, the R^2 for the other equations show that the ABILITY score does not really contribute to the variation in the parameters.

The equivalent LISREL syntax is given in **trial2b.lis**.