Homework Assignment for Lesson 6

*Instructions: Please submit a single document (or pdf file) to the ANGEL dropbox for Lesson 6. You need not complete these analyses with graphs or diagnostics. Just focus on specific questions asked.*

1) (25 pts) A hospital is interested in evaluating inter-rater reliability in nurse assessment of dementia. A random selection of 4 nurses evaluated 6 patients (randomly assigned) and recorded a dementia score.

Dataset = HW6\_Nurse\_ICC\_data

Use a one-way, random effects model (software of your choice) to find

1) the variance components,

|  |  |
| --- | --- |
| Minitab | SAS |
| The output we get:  **General Linear Model: rating versus nurse**  Factor Information  Factor Type Levels Values  nurse Random 4 1, 2, 3, 4  Analysis of Variance  Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value  nurse 3 97.46 57.68% 97.46 32.486 9.09 0.001  Error 20 71.50 42.32% 71.50 3.575  Total 23 168.96 100.00%  Model Summary  S R-sq R-sq(adj) PRESS R-sq(pred)  1.89077 57.68% 51.33% 102.96 39.06%  Coefficients  Term Coef SE Coef 95% CI T-Value P-Value  Constant 5.292 0.386 ( 4.487, 6.097) 13.71 0.000  nurse  1 2.375 0.668 ( 0.981, 3.769) 3.55 0.002  2 -2.792 0.668 (-4.186, -1.397) -4.18 0.000  3 -0.958 0.668 (-2.353, 0.436) -1.43 0.167  Regression Equation  rating = 5.292 + 2.375 nurse\_1 - 2.792 nurse\_2 - 0.958 nurse\_3 + 1.375 nurse\_4  Equation treats random terms as though they are fixed.  Fits and Diagnostics for Unusual Observations  Obs rating Fit SE Fit 95% CI Resid Std Resid Del Resid HI Cook’s D  8 2.000 6.667 0.772 (5.057, 8.277) -4.667 -2.70 -3.31 0.166667 0.37  Obs DFITS  8 -1.47952 R  R Large residual  Expected Mean Squares, using Adjusted SS  Expected Mean Square  Source for Each Term  1 nurse (2) + 6.0000 (1)  2 Error (2)  Error Terms for Tests, using Adjusted SS  Synthesis  Source Error DF Error MS of Error MS  1 nurse 20.00 3.5750 (2)  Variance Components, using Adjusted SS  Source Variance % of Total StDev % of Total  nurse 4.81852 57.41% 2.19511 75.77%  Error 3.575 42.59% 1.89077 65.26%  Total 8.39352 2.89716 |  |

Therefore the variance components are:

Source Variance % of Total

nurse 4.81852 57.41%

Error 3.575 42.59%

Total 8.39352

2) calculate the Intraclass Correlation Coefficient (ICC)

ICC is computed as:

ICC = s2among trts / (s2among trts + s2within trts)

ICC = 4.81852 / (4.81852 + 3.575) = 0.5741

Note that there are many variants of ICC calculations, and we are simply asking here for a ‘between vs. within’ partition. So you won’t be including ‘patient’ in your model. Patient-to-patient variability and nurse x patient interaction will be included in the error (‘within’) term.

2) (75 pts) A survey of health awareness in households (an index) is made in 3 states, 3 cities within each state, and sampling 5 households within each of cities.

The dataset is HW6\_Health\_data.

1. With software of choice you will run three separate analyses. First, run a Model I (all factors as fixed effects). Second, run a Model II (all factors as random effects). Finally run a Model III ANOVA (a mixed model with state as a fixed effect).

**Model I – Treat all factors as fixed and nested.**

|  |  |
| --- | --- |
| Minitab | SAS |
| Factor Information  Factor Type Levels Values  state Fixed 3 1, 2, 3  city(state) Fixed 9 1(1), 2(1), 3(1), 1(2), 2(2), 3(2), 1(3), 2(3), 3(3)  Analysis of Variance  Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value  state 2 6976.8 60.28% 6976.8 3488.4 32.26 0.000  city(state) 6 704.3 6.08% 704.3 117.4 1.09 0.390  Error 36 3893.2 33.64% 3893.2 108.1  Total 44 11574.3 100.00% |  |

**Model II - All factors as random and nested**

|  |  |
| --- | --- |
| Minitab | SAS |
| Factor Information  Factor Type Levels Values  state Random 3 1, 2, 3  city(state) Random 9 1(1), 2(1), 3(1), 1(2), 2(2), 3(2), 1(3), 2(3), 3(3)  Analysis of Variance  Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value  state 2 6976.8 60.28% 6976.8 3488.4 29.72 0.001  city(state) 6 704.3 6.08% 704.3 117.4 1.09 0.390  Error 36 3893.2 33.64% 3893.2 108.1  Total 44 11574.3 100.00%  Model Summary  S R-sq R-sq(adj) PRESS R-sq(pred)  10.3993 66.36% 58.89% 6083.13 47.44%  Variance Components, using Adjusted SS  Source Variance % of Total StDev % of Total  state 224.736 67.14% 14.9912 81.94%  city(state) 1.84667 0.55% 1.3589 7.43%  Error 108.144 32.31% 10.3993 56.84%  Total 334.727 18.2956 |  |

**Model III - State as fixed and City as random and nested**

|  |  |
| --- | --- |
| Minitab | SAS |
| Factor Information  Factor Type Levels Values  state Fixed 3 1, 2, 3  city(state) Random 9 1(1), 2(1), 3(1), 1(2), 2(2), 3(2), 1(3), 2(3), 3(3)  Analysis of Variance  Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value  state 2 6976.8 60.28% 6976.8 3488.4 29.72 0.001  city(state) 6 704.3 6.08% 704.3 117.4 1.09 0.390  Error 36 3893.2 33.64% 3893.2 108.1  Total 44 11574.3 100.00%  Variance Components, using Adjusted SS  Source Variance % of Total StDev % of Total  city(state) 1.84667 1.68% 1.3589 12.96%  Error 108.144 98.32% 10.3993 99.16%  Total 109.991 10.4877 |  |

1. For each model, state the Null hypotheses for each effect.

**Model I – Treat all factors as fixed and nested.**

State:

H0: αState1 = α State2 = α State3 = 0 or equivalently H0: all αi = 0 vs. HA = not all βj(i) = 0

City: Nested within State

H0 = all βj(i) = 0 vs. HA = not all βj(i) = 0

**Model II - All factors as random and nested**

State:

H0: σα2 = 0 vs. HA: σα2 > 0

City: Nested within State

H0: σβj(i)2 = 0 vs. HA: σβj(i)2 > 0

**Model III - State as fixed and City as random and nested**

State:

H0: αState1 = α State2 = α State3 = 0 or equivalently H0: all αi = 0 vs. HA = not all βj(i) = 0

City: Nested within State

H0: σβj(i)2 = 0 vs. HA: σβj(i)2 > 0

c) Include the ANOVA tables from each model and summarize how the conclusions change depending on the model.

**Model I – Treat all factors as fixed and nested.**

|  |  |
| --- | --- |
| Minitab | SAS |
| Factor Information  Factor Type Levels Values  state Fixed 3 1, 2, 3  city(state) Fixed 9 1(1), 2(1), 3(1), 1(2), 2(2), 3(2), 1(3), 2(3), 3(3)  Analysis of Variance  Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value  state 2 6976.8 60.28% 6976.8 3488.4 32.26 0.000  city(state) 6 704.3 6.08% 704.3 117.4 1.09 0.390  Error 36 3893.2 33.64% 3893.2 108.1  Total 44 11574.3 100.00% |  |

Summary:

* State is significant. P-value is < 0.001
* City(state) is not at alpha=5% (0.05 assumed). P-value is < 0.390

**Model II - All factors as random and nested**

|  |  |
| --- | --- |
| Minitab | SAS |
| Factor Information  Factor Type Levels Values  state Random 3 1, 2, 3  city(state) Random 9 1(1), 2(1), 3(1), 1(2), 2(2), 3(2), 1(3), 2(3), 3(3)  Analysis of Variance  Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value  state 2 6976.8 60.28% 6976.8 3488.4 29.72 0.001  city(state) 6 704.3 6.08% 704.3 117.4 1.09 0.390  Error 36 3893.2 33.64% 3893.2 108.1  Total 44 11574.3 100.00%  Model Summary  S R-sq R-sq(adj) PRESS R-sq(pred)  10.3993 66.36% 58.89% 6083.13 47.44%  Variance Components, using Adjusted SS  **Source Variance % of Total**  state 224.736 67.14%  city(state) 1.84667 0.55%  Error 108.144 32.31%  Total 334.727 |  |

Summary:

* This is a fully nested, random effects model.
* State is significant
* City(state) is not significant at alpha=5% (0.05 assumed). P-value is < 0.390
* The largest (and significant) variance components are the state, and the house-to-house variation within the state and city (the residual).
* The city(state) variance components is negligible (insignificant)

**Model III - State as fixed and City as random and nested**

|  |  |
| --- | --- |
| Minitab | SAS |
| Factor Information  Factor Type Levels Values  state Fixed 3 1, 2, 3  city(state) Random 9 1(1), 2(1), 3(1), 1(2), 2(2), 3(2), 1(3), 2(3), 3(3)  Analysis of Variance  Source DF Seq SS Contribution Adj SS Adj MS F-Value P-Value  state 2 6976.8 60.28% 6976.8 3488.4 29.72 0.001  city(state) 6 704.3 6.08% 704.3 117.4 1.09 0.390  Error 36 3893.2 33.64% 3893.2 108.1  Total 44 11574.3 100.00%  Variance Components, using Adjusted SS  Source Variance % of Total StDev % of Total  city(state) 1.84667 1.68% 1.3589 12.96%  Error 108.144 98.32% 10.3993 99.16%  Total 109.991 10.4877 | The results for hypothesis tests for the fixed effects appear as:    and a hypothesis test for the random effect appears as: |

Summary:

* The last part of the SAS output shown above is produced by the “covtest” option in the proc mixed statement and indicates that the H0:σ2city(state)=0 cannot be rejected α = 0.05 (the p-value is 0.8985). It is important to point out here that this result is consistent with the result that we see in the ANOVA table above only in the outcome. The p-values are different but both have city(state) as insignificant (P > F = 0.3896).
* State is significant
* City(state) is not significant at alpha=5% (0.05 assumed). P-value is < 0.390

d) Given that researchers want to have a nation-wide scope of inference, how do you recommend that they work with the state factor in designing this experiment?

Since the scope of interest is the entire nation, it will be best to treat both state and city as random factors.