Math 158 – Spring 2009
Jo Hardin
warm-up # 9

Name: ________________________________

The data below are intelligence test scores for 38 adopted children categorized according to whether they child's biological & adopted parents fell into the highest or lowest socio-economic status (SES). (Low SES means being poor, high SES means being wealthy.) – SES I: both sets of parents were in low SES; SES II: biologic parents were in high SES, adopted parents were in low SES; SES III: biologic parents were in low SES, adopted parents were in high SES; SES IV: both sets of parents were in high SES.

group	sample size	mean	SD
SES I	10	92.4	15.41
SES II	8	107.5	11.94
SES III	10	103.6	12.71
SES IV	10	119.6	12.24

1. By writing out sums of squares for a full and reduced model, test whether environment is a factor in IQ. Specifically, test:

$$\begin{aligned} H_0 :& \quad \mu_I = \mu_{II} \text{ and } \mu_{III} = \mu_{IV} \\ H_a :& \quad \text{not } H_0 \end{aligned}$$

   (This is an F-test, remember to write down your hypotheses, compute a test statistic, find the p-value, and interpret your results in the terms of the problem.) Hint: you'll want to calculate SSTR for the reduced model (instead of trying to calculate SSE directly for the reduced model).

2. From this data test the hypothesis that there is at least one group difference between these 4 SES classes.

3. Suppose there is interest in the effect of biological parents socio-economic status on the distribution of intelligence scores. We may investigate this by considering $\mu_4 - \mu_3$ (if we consider only children whose adoptive parents are in the high SES) and by $\mu_2 - \mu_1$ (if we consider only children whose adoptive parents are in the low SES.) We might also look at the average of these two quantities: $\frac{1}{2}(\mu_4 - \mu_3) + \frac{1}{2}(\mu_2 - \mu_1)$. Find an estimate and a 95% CI for this amount. Interpret the CI in the words of the problem.

4. Are we able to claim that being in a particular SES causes high (or low) intelligence scores? Should we worry about confounding variables? Explain, and give possible reasons for your explanation.

Solution:

1. From the full model:
$MSE = s_p^2 = \frac{9(15.41)^2+7(11.94)^2+9(12.71)^2+9(12.24)^2}{34} = \frac{5937.41}{34} = 174.63, s_p = 13.21$
$SSE = MSE \cdot 34 = 5937.41$
$\overline{Y}_{..} = 10 * 92.4 + 8 * 107.5 + 10 * 103.6 + 10 * 119.638 = 105.68$
SSTR $= 10(105.68 - 92.4)^2 + 8(105.68 - 107.5)^2 + 10(105.68 - 103.6)^2 + 10(105.68 - 119.6)^2 = 3771$
MSTR $= \frac{\text{SSTR}}{df} = \frac{3771}{3} = 1257$
$SSTO = SSE + SSTR = 5937.41 + 1257 = 9708.4$

   From the reduced model:
   If we let low be SES I and SES II and high be SES III and SES IV, then
   $\overline{Y}_{low} = \frac{10*92.4+8*107.5}{18} = 99.11, \overline{Y}_{high} = \frac{103.6+119.6}{2} = 111.6$
   SSTR(R) $= 18 * (99.11 - 105.68)^2 + 20 * (111.6 - 105.68)^2 = 1477.8962$
   SSE(R) = SSTO - SSTR(R) = 9708.4 - 1477.9 = 8230.5

   $F^* = \frac{(SSE(R)-SSE(F))/(df(R)-df(F))}{MSE(F)} = \frac{(8230.5-5937.41)/2}{174.63} = 6.57$
   $p - value = P(F_{2,34} \geq 6.57) = 1 - pf(6.57, 2, 34) = 0.0039$

   We do reject the null hypothesis. Once we've accounted for the difference between the environments, there is still some significant difference within the remaining groups (that is, there is some remaining significant difference due to biology).

2. $F^* = \frac{\text{MS(btwn)}}{s_p^2} = \frac{727.93}{174.63} = 4.168$
df= (numerator = 3, denominator = 34), p-value ≈ 0.01
The small p-value gives an indication that the means are not all equivalent across the four groups. There is a difference in intelligence scores across the SES groups.

source	SS	df	MS	F	p-val
between	3771	3	1257	7.2	0.0007
within	5937.42	34	174.63	X	X
total	9708.4	37	X	X	X

3. the point estimate is: $\frac{1}{2}(119.6 - 103.6) + \frac{1}{2}(107.5 - 92.4) = 15.55$
the SE of this estimate is: $13.21\sqrt{\frac{(1/2)^2}{10} + \frac{(1/2)^2}{8} + \frac{(1/2)^2}{10} + \frac{(1/2)^2}{10}} = 2.95$
multiplier is t (one planned comparison): df=34, 2.042
95% CI is: $15.55 \pm 2.042 \times 2.95$, (9.52, 21.58)
There is a biological parent effect, the groups whose biologic parents are in a high SES have, on average, 9.52 to 21.58 points more on the intelligence test than those whose biologic parents are in the low SES.

4. Yes, we worry, this is not a randomized study. What confounding variables can you think of?