Psychology 407
Assignment E

Suppose we randomly assign 60 people to three groups and ask each member in the group to rate a specific geometric figure on a 1 to 10 point scale based on its "complexity". A large rating will imply a more complex figure, and all members in a group rate the same figure independently. The data and figures rated by each group are as follows.

Group 1	Group 2	Group 3
6	10	5
2	8	10
2	9	10
6	4	3
3	5	10
1	3	10
4	3	10
4	4	10
2	5	10
4	2	7
4	10	10
3	3	10
6	8	4
1	7	10
2	9	9
5	7	1
2	8	10
2	9	10
2	6	10
2	9	9

Analyze these data using the one-way (single-factor) fixed-effects analysis-of-variance model, i.e.,

$$\text{(I) } y_{ij} = \mu_i + \epsilon_{ij} \ ,$$

where $\epsilon_{ij} \sim N(0, \sigma^2)$ and are independent; Y_{ij} is the value for the response variable for person j within group i; μ_i are fixed parameters; $1 \leq i \leq r$ (the number of groups); $1 \leq j \leq n_i$ (the number of observations in group i).

a) Construct the appropriate analysis-of-variance (ANOVA) table.

b) What are the estimates of μ_i, $1 \leq i \leq r$; and of σ^2?

c) Show (no need to compute anything) schematically how the analysis-of-variance model given in (I) can be recast in matrix terms as a linear model having the usual form: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{X}$ contains 3 indicator variables for its columns. Indicate the size of the matrices and what they contain.

Assuming that the model given in (c) if the Full model, what is the form of the Reduced model that would be used in testing $H_0 : \mu_1 = \mu_2 = \mu_3$? Considering the general test statistic in Kutner, et al. (p. 75, equation 2.70), what are the numerical values for its constituent terms (Hint: use the numerical results in (a))?

Consider the "effect" form of the one-way ANOVA model:

$$\text{(II)}\ y_{ij} = \mu. + \tau_i + \epsilon_{ij}\ ,$$

where $\mu. = \Sigma_{i=1}^{r} n_i \mu_i / n_T$, and $n_T = \Sigma_{i=1}^{r} n_i$.

e) What are the estimates of τ_i, $1 \leq i \leq r$? In general, would these be different if we define $\mu. = \Sigma_{i=1}^{r} \mu_i / r$? In our example?

f) In defining the general linear model for the form in (II), what would $\mathbf{X}$ look like if

$$\boldsymbol{\beta} = \begin{bmatrix} \mu. \\ \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} ?$$

Why is it "impossible" to use a $\boldsymbol{\beta}$ of this form computationally in getting an estimate of $\boldsymbol{\beta}$?

g) If $\boldsymbol{\beta}$ is given as

$$\begin{bmatrix} \mu_{\cdot} \\ \tau_1 \\ \tau_2 \end{bmatrix},$$

what would $\mathbf{X}$ look like in general if $\mu_{\cdot} = \sum_{i=1}^{r} n_i \mu_i / n_T$ or $\mu_{\cdot} = \sum_{i=1}^{r} \mu_i / r$. In our example, are there any differences? Why?

h) Assuming that the "true" means are $\mu_1 = 4$, $\mu_2 = 4$, and $\mu_3 = 7$ and $\sigma^2 = 6$, what is the power for rejecting $H_0 : \mu_1 = \mu_2 = \mu_3$ given our sample sizes (and $\alpha = .05$)?