Unit 6 – Analysis of Variance
Practice Problems (1 of 2)

Due: Wednesday November 8, 2023

Last date to submit late for credit (-20 points): Wednesday November 15, 2023

Before you begin. Download from the course website
anova_infants.xlsx

Zelazo et al. (1972) studied the variability in age at first walking in infants. 24 infants were randomly assigned to four groups of equal sample size (6 infants per group), with groups defined by method of reinforcement of walking: (1) active (2) passive (3) no exercise; and (4) 8 week control. The outcome variable measured was age at first walking, in months. The following table lists the study data, by group.

Table – Study Data of Zelazo et al (1972), n=24:

Active Group	Passive Group	No-Exercise Group	8 Week Control
9.00	11.00	11.50	13.25
9.50	10.00	12.00	11.50
9.75	10.00	9.00	12.00
10.00	11.75	11.50	13.50
13.00	10.50	13.25	11.50
9.50	15.00	13.00	12.35

Source: Zelazo et al (1972) "Walking" in the newborn. *Science* 176: 314-315.

Data dictionary/Codebook:

Variable	Label	Type	Coding
group	Group	numeric	1 = active 2 = passive 3 = noex 4 = control
age	Age, months	numeric	Continuous, months
I_active	Indicator group = "active"	numeric	1 if group = 1 ("active") 0 otherwise
I_passive	Indicator group="passive"	numeric	1 if group = 2 ("passive") 0 otherwise
I_noex	Indicator group = "noex"	numeric	1 if group = 3 ("noex") 0 otherwise

#1.

Deviation from means. State the analysis of variance model using deviation from means notation μ and τ_i and σ^2 as appropriate. Define all terms and constraints on the parameters.

#2.

By any means you like, produce a side by side box plot showing the distribution of age at first walking, separately for each of the 4 groups.

#3.

By any means you like, obtain the entries of the analysis of variance table for this one way analysis of variance. Use your computer output (or excel work or hand calculations or whatever) to complete the following table:

Source	df	Sum of Squares SSQ	Mean Square MSQ	F-Statistic	p-value
Between Groups					
Within Groups					
Total, corrected					

#4.

Write a 2-5 sentence report of your description and hypothesis test findings using language as appropriate for a client who is intelligent but is not knowledgeable about statistics. Consider including a figure and/or table that you think is appropriate.

#5.

Reference cell coding Repeat your analysis, this time using what you learned in Unit 5 - Normal Theory Regression. Specifically, using appropriately defined indicator variables, perform a multivariable linear regression analysis of these same data! Use your computer output to complete the following table:

Source	df	Sum of Squares SSQ	Mean Square MSQ	F-Statistic	p-value
Due Model					
Due Error (residual)					
Total, corrected					

#6.

Deviation from means *and* ***reference cell coding*** **are equivalent!** Using your output from your two analyses (1st-analysis of variance, 2nd – regression), obtain the predicted mean of Y =age at first walking twice in two ways.

	Prediction Using One Way Analysis of Variance	**Prediction Using Multiple Linear Regression**
Active		
Passive		
No-Exercise		
Control		